이학박사학위논문

# 마이크로RNA 생성효소인 Dicer의 기질 인지 기전에 대한 연구

# Study on the mechanism of substrate recognition by Dicer

2015년 8월

서울대학교 대학원

생명과학부

박종은

# Study on the mechanism of substrate recognition by Dicer

**by**

**Jong-eun Park**

**Advisor**

**Professor V. Narry Kim, D. Phil.**

**A Thesis for the Degree of Doctor of Philosophy**

**August, 2015**

**School of Biological Sciences**

**Seoul National University**

# ABSTRACT

MicroRNAs (miRNAs) are small noncoding RNAs which mediate post-transcriptional gene silencing by base-pairing to target mRNAs. Deep sequencing data for small RNAs indicate that mature miRNAs have sequence variations at the 3′ ends, which arise mainly from untemplated A or U additions (1-2 nt). The biological significance of these variations is under extensive investigation. The 3′ end modification is not confined to mature miRNAs. Sequence analysis of pre-miRNAs revealed that some pre-miRNAs also have 3′ end variations, mostly untemplated U additions. According to the previous model for processing of pre-miRNAs, Dicer selects the cleavage sites by measuring a set distance from the 3′ overhang of the pre-miRNA terminus. This model predicts that upon 3′ modification of pre-miRNA, Dicer cleavage site would be shifted, changing the sequence and targeting potential of mature miRNA. To determine the effects of the pre-miRNA end variation on Dicer processing, I performed in vitro processing assays using pre-miRNAs with varying length of 3′ overhang. I found that Dicer cleavage site remains unaltered despite the different length of the 3′ overhang. Further analysis indicates that human Dicer anchors not only the 3′ end but also the 5′ end, with the cleavage site determined mainly by the distance (~22 nt) from the 5′ end (5′ counting rule). This mode of cleavage requires a 5′-terminal

phosphate group and is facilitated by unstable base-pairing of duplex terminus. In a physiological concentration of magnesium ion, 5′ end-dependent cleavage is favored, which suggests that the 5′ counting rule would prevail in vivo. I further identify a novel basic motif (5′-pocket) in human Dicer, for the recognition of the 5′-phosphorylated end. The 5′ counting rule and the 5′-anchoring residues are conserved in miRNA-generating enzymes such as Drosophila Dicer-1 but not in Giardia Dicer.

Mutations in the 5′-pocket reduce processing efficiency and alter cleavage sites in vitro. Consistently, miRNA biogenesis is perturbed in vivo when Dicer-null embryonic stem cells are replenished with the 5′-pocket mutant. Thus, the 5′-phosphorylated end recognition by Dicer is important for precise and effective biogenesis of miRNAs. Also, in collaboration with Dr. Dinshaw Patel group, the structure counterpart of 5′-pocket was discovered in the crystal structure of Dicer PAZ cassette and siRNA duplex, providing structural evidence for pre-miRNA recognition model by Dicer. Insights from this study should also afford practical benefits to the design of small hairpin RNAs.

**Key words**

microRNA (miRNA); miRNA biogenesis; Dicer; PAZ domain; pre-microRNA (pre-miRNA); uridylation; processing; small RNA; sequencing; crystal structure; evolution; RNase III; gene regulation; RNA silencing

# CONTENTS

# LIST OF FIGURES AND TABLES

# ABBREVIATIONS

| | |
|---|---|
| **Ago,** | Argonaute |
| **ATP,** | adenosine triphosphate |
| **bp,** | base pair |
| **DGCR8,** | DiGeorge syndrome critical region 8 |
| **DNA,** | deoxyribonucleic acid |
| **dsRNA,** | double stranded RNA |
| **ES cells**, | embryonic stem cells |
| **EXP5,** | Exportin 5 |
| **FDR,** | false discovery rate |
| **GEO,** | gene expression omnibus |
| **miRNA,** | microRNA |
| **mRNA,** | messenger RNA |
| **nt,** | nucleotide |
| **PAZ,** | Piwi Argonaute Zwille |
| **piRNA,** | Piwi-interacting RNA |
| **pre-miRNA,** | precursor microRNA |
| **pri-miRNA,** | primary microRNA |
| **qRT-PCR,** | quantitative real-time polymerase chain reaction |
| **RISC,** | RNA induced silencing complex |

| | |
|---|---|
| **RITS,** | RNA-induced transcriptional silencing |
| **RNA,** | ribonucleic acid |
| **RNase III,** | ribonuclease III |
| **RPM**, | read per million |
| **RT,** | reverse transcription |
| **siRNA,** | small interfering RNA |
| **ssRNA,** | single stranded RNA |
| **TRBP,** | TAR RNA-binding protein |
| **U-tail,** | uridine tail |
| **UTP,** | uridine triphosphate |
| **UTR,** | untranslated region |

# CHAPTER I

# Introduction

## I-1. Small RNA pathways

Since the first discovery of small RNA lin-4 in C. elegans (Lee et al., 1993), diverse class of small RNAs have been found in eukaryotic organisms (Kim et al., 2009). Eukaryotic small RNAs share limited size range of 20-30 nts and exert their role in association with Argonaute (Ago)-family proteins. According to their biogenesis mechanism and type of interacting Ago proteins, small RNAs can be classified into three main classes: microRNAs (miRNAs), small interfering RNAs (siRNAs), and Piwi-interacting RNAs (piRNAs). Although most small RNAs play suppressive roles in gene expression, each small RNA pathway utilizes distinct action mechanisms, ranging from heterochromatin formation in the nucleus to mRNA destabilization and translation control in the cytoplasm (Hammond et al., 2001; Olsen and Ambros, 1999; Volpe et al., 2002).

Among small RNAs found in humans, miRNAs have been under the most extensive investigation due to their universal expression and diverse regulatory functions. Animal miRNAs are single-stranded RNAs (ssRNAs) which are ~22 nt in length and generated from endogenous hairpin-shaped transcripts (Bartel, 2004). MiRNAs recruit the Ago protein to target mRNA through complementary base paring between their seed sequence (positions 2-7 nt relative to the 5′ end of miRNA) and the 3′ untranslated region (UTR), which then represses the expression of target mRNA through deadenylation

medicated decay or translational inhibition (Bartel, 2009). Target of miRNAs has been discovered through genetic, biochemical and computational approaches, which revealed that at least one third of all human genes are regulated by miRNAs (Lewis et al., 2005). Moreover, function of miRNAs have been implicated in various biological pathways including stem cell differentiation, tissue development, and tumorigenesis (Ambros, 2004). Considering their biological importance and large number of targets, miRNAs are being considered as not only potent target of drug development but also platform for gene expression engineering.

## I-2. Biogenesis of MicroRNA

Animal miRNAs are transcribed in the nucleus by RNA polymerase II to yield a nascent precursor form named as primary miRNAs (pri-miRNAs). MiRNA sequences are embedded in the hairpin structure located in this pri-miRNAs, which are then successively cleaved by two RNase III enzymes, Drosha and Dicer. Firstly, microprocessor complex which consists of Drosha and DiGeorge syndrome critical region 8 (DGCR8) recognizes the hairpin structure and cleaves basal stem region to release ~70 nt hairpin-shaped precursor miRNAs (pre-miRNAs) (Denli et al., 2004; Gregory et al., 2004; Han et al., 2004; Han et al., 2006; Landthaler et al., 2004; Lee et al., 2003). Pre-miRNAs are then exported to the cytoplasm by interacting with exportin-

5 (EXP5) (Bohnsack et al., 2004; Lund et al., 2004; Yi et al., 2003), recognized by Dicer and its cofactor transactivation response RNA-binding protein (TRBP) and processed into mature miRNA duplex of ~22 nucleotides (Grishok et al., 2001; Hutvagner et al., 2001; Ketting et al., 2001; Knight and Bass, 2001). Finally, the duplex miRNAs is incorporated to the Ago protein, which is soon followed by expulsion of one strand, resulting in functional RNA induced silencing complex (RISC). For the selection of guide strand, thermodynamic stability of two ends of duplex plays an important role (Hammond et al., 2001; Mourelatos et al., 2002; Tabara et al., 1999).

As mature miRNA sequence is determined by Drosha and Dicer cleavage, accurate substrate recognition and cleavage site selection of these enzymes are important. Proteins involved in miRNA biogenesis pathway developed multiple mechanisms to ensure that only miRNAs are incorporated into functional RISC. One of such mechanisms is mediated by PAZ (Piwi Argonaute Zwille) domain in the Dicer and Ago proteins. Biochemical study on the function of PAZ domains revealed that they have RNA-binding surface with a specificity to the 3′ overhang (Lingel et al., 2004; Ma et al., 2004; Song et al., 2003; Yan et al., 2003), which is the characteristic shared by pre-miRNAs and mature miRNA duplexes. Thus, PAZ domain acts as a keyhole which allows only the RNAs bearing the 3′ overhang as Drosha-generated key

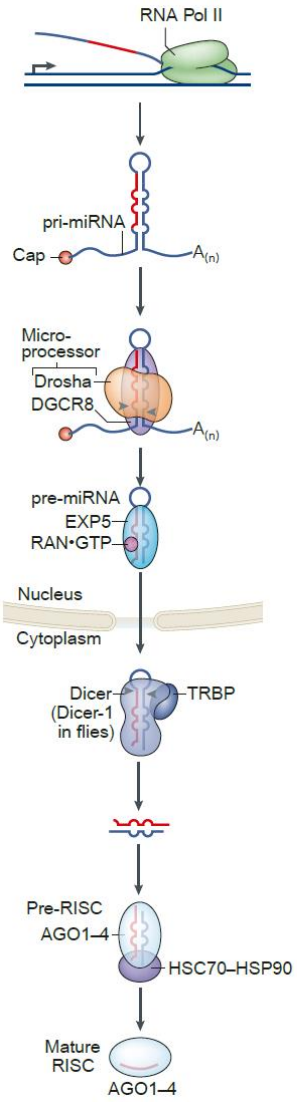



**Figure I-1. Canonical miRNA biogenesis pathway[1]**

[1] This figure is modified from: Ha, M. and V. N. Kim (2014). Regulation of microRNA biogenesis. Nat Rev Mol Cell Biol 15(8): 509-524.

to pass through the Dicer and Ago gate. However, as there are many other RNAs such as tRNAs which have similar 3′ overhang structures, actual substrate recognition mechanisms for the Dicer and Ago proteins are much more complex. Moreover, as small RNA pathways are highly diversified among different species, proteins are also highly evolved to meet the specific requirement in respective pathways.

## I-3. Characteristics of Dicer

The Dicer proteins belong to RNase III enzyme family and are broadly conserved in eukaryotic organisms. Dicer is key player in the miRNA biogenesis pathway, which cleaves dsRNA within helical segments with site selectivity. Dicer functions as a molecular ruler, generating products of defined length (21-25 nt depending on the homologue and species) with a 5′-phosphate and a 2 nt overhang at the 3′ end (Elbashir et al., 2001; Provost et al., 2002; Zhang et al., 2002). Human Dicer is a large protein of ~200 kDa which contains, from N terminus to C terminus, a DEAD/helicase domain, a dsRNA-binding fold domain, a PAZ domain, two RNase III domains and a double-stranded RNA-binding domain. While bacterial RNase III protein has one RNase III domain and function as homodimer (Blaszczyk et al., 2001), two RNase III domain of human Dicer form a intramolecular pseudo-dimer

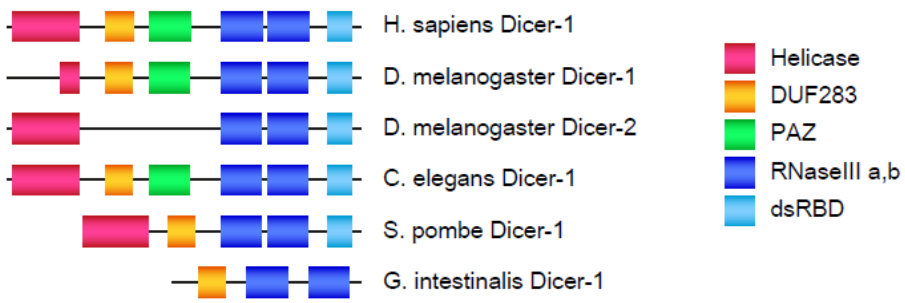



**Figure I-2. Domain structure of Dicer proteins**

which constitutes a single-processing center containing two catalytic sites (Zhang et al., 2004). Each catalytic site is positioned to cleave each strand with 2-nt interval, generating canonical a 2 nt 3′ overhang structure on the products. The PAZ domain, which has binding affinity towards the 3′ protrusion of RNA molecule, anchors Dicer to the end of dsRNA, from which Dicer measures defined distance to locate the cleavage site (MacRae et al., 2007; Macrae et al., 2006). Role of the DEAD/helicase domain has been controversial. In Drosophila, siRNA processing enzyme Dicer-2 contains ATPase activity in helicase domain while miRNA processing Dicer-1 does not have the activity (Lee et al., 2004; Liu et al., 2003). Also, Dicer-2 requires ATP hydrolysis for processing of long dsRNA, suggesting that helicase domain facilitates processive cleavage of Dicer (Cenik et al., 2011; Welker et al., 2011). However, human Dicer-1 which mainly generates miRNAs by cutting pre-miRNA once still contains functional helicase domain and helicase domain of human Dicer has been shown to confer specificity to pre-miRNAs (Gu et al., 2012; Ma et al., 2008; Tsutsumi et al., 2011), implying that helicase domain might have different functions in different organisms.

Basic domain structure of Dicer can be found in organisms representing all eukaryotic lineages, implying that Dicer was invented at a very early stage of eukaryotic evolution. However, as small RNA pathways diverged among species, there may have been concomitant variations on the domain structure of Dicer (Cerutti and Casas-Mollano, 2006; Mukherjee et al., 2013; Shabalina

and Koonin, 2008). In most extreme cases such as budding yeast or parasitic protozoa, Dicer genes disappeared along with the loss of RNA interference pathway (Baum et al., 2009; Drinnenberg et al., 2009; Ullu et al., 2004). In fission yeast, Dicer generates siRNAs which are incorporated to RNA-induced transcriptional silencing (RITS) complex to induce heterochromatin formation (Volpe et al., 2002). Interestingly, fission yeast Dicer does not recognizes the end of dsRNA, instead, this enzyme induce internal cleavage of substrate (Colmenares et al., 2007). Correspondingly, sequence analysis of fission yeast Dicer revealed that this enzyme has lost PAZ domain, which is responsible for RNA end recognition. In some animal species, Dicer gene has been duplicated to yield Dicer-1 and Dicer-2 (Lee et al., 2004). In general, the Dicer-1 group is more specialized for hairpin-shaped pre-miRNAs while the Dicer-2 group plays a role in antiviral defense. Accordingly, the Dicer-1 group repeatedly lost the ATPase function of helicase domain, which has been implicated in processive cleavage and/or recognition of 5′ overhang/blunt end of dsRNAs. In plants, there are four Dicer-like proteins with distinct roles: Dcl-1 is involved in miRNA production; Dcl-2 is involved in antiviral defense; Dcl-3 generates siRNAs which induce chromatin modification and transcriptional silencing; and Dcl-4 produces siRNAs to regulate gene expression (Borsani et al., 2005; Gasciolli et al., 2005; Kurihara and Watanabe, 2004; Park et al., 2002; Vazquez et al., 2004; Xie et al., 2005; Xie et al., 2004).

## I-4. End recognition mechanism of Dicer

As I mentioned above, precise selection of cleavage sites by RNase III enzymes is critical in miRNA biogenesis because alterations in the cleavage site can change the abundance and/or targeting specificity of the miRNA. To determine the cleavage site, Drosha and Dicer recognize certain RNA structures and cleave a fixed distance away from the structure. In the case of Drosha, it binds to the base of the stem-loop and locates the catalytic site of Drosha ~11 bp away from the ssRNA-dsRNA junction (Han et al., 2006). Thus, the ssRNA-dsRNA junction serves as the reference point for Drosha processing.

Dicer, on the other hand, is known to measure ~22 nt away from the 3′ end of the open terminus of dsRNA helices (Vermeulen et al., 2005; Zhang et al., 2002; Zhang et al., 2004). Like Drosha, Dicer has two tandem RNase III domains (RIIIDs) that form an intramolecular dimer to compose a processing center. A typical Dicer homologue also contains an N-terminal DExD/H helicase-like domain, a DUF283 domain, a PAZ domain, and a C-terminal dsRNA binding domain (dsRBD). Because the PAZ domain is found in Dicer as well as in Ago, and PAZ from Ago interacts with the 3′ terminus of RNA (Lingel et al., 2004; Ma et al., 2004; Song et al., 2003; Yan et al., 2003), it was proposed that Dicer may also bind to the 3′ end through its PAZ domain.

The crystal structure of Dicer from Giardia intestinalis in the free state showed that the processing center is connected to the PAZ domain via a flat, positively charged extension (Macrae et al., 2006). The distance (65 Å) between the 3′ end-binding pocket and the catalytic center corresponds to the length of the product generated by Giardia Dicer (~25 nt). A truncated mutant of Giardia Dicer lacking the PAZ domain yielded RNA products of heterogeneous lengths (MacRae et al., 2007). Based on these observations, it was proposed that the PAZ domain anchors the 3′ overhang of an open terminus, and the dsRNA stem is placed along the positively charged protein extension to reach the catalytic center of Dicer (MacRae et al., 2007). This spatial arrangement would enable Dicer to measure a fixed distance from the 3′ end of the terminus, which is consistent with previous biochemical evidences from human Dicer (Vermeulen et al., 2005; Zhang et al., 2004). Here, I refer to this model as the "3′ counting model."

Recent studies have shown that certain pre-miRNAs are modified at their 3′ end in the cell. The most common type of pre-miRNA modification is addition of non-templated uridyl residues (Burroughs et al., 2010; Chiang et al., 2010; Heo et al., 2008; Wu et al., 2009). The uridylation of pre-miRNA has been also validated by our group (**Figure I-3**). In the case of pre-miR-24-2, more than 50% had U addition in mouse embryonic stem cells. According to the current 3′ counting model, such 3′ end modifications are expected to result in a shift of the Dicer cleavage site towards the open terminus. This

would change the seed sequences of miRNAs originating from the 3′ strand and may even alter strand selection (Bartel, 2009; Tomari et al., 2004).

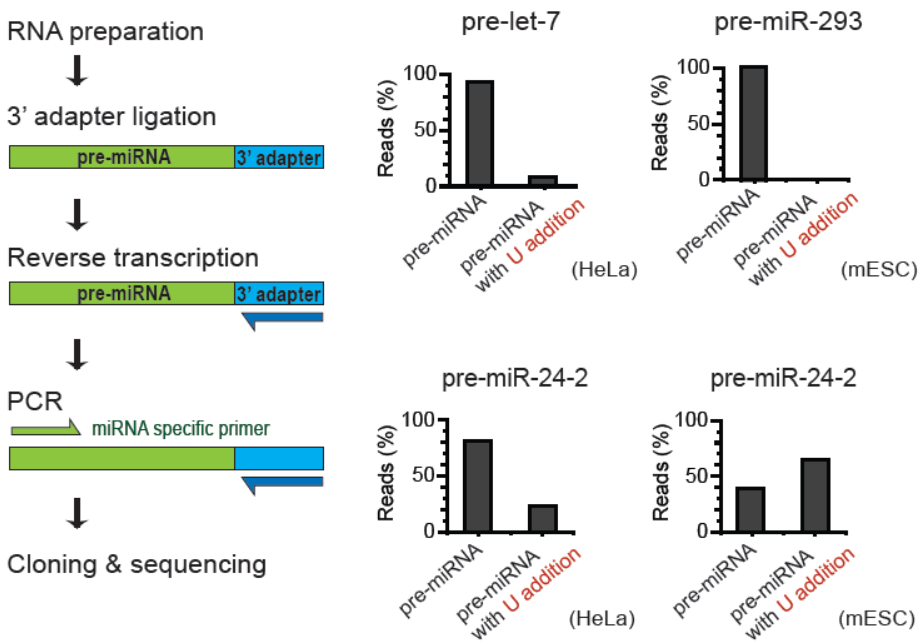



**Figure I-3 Frequent uridylation of pre-miRNAs**[2]

To sequence 3′ end of pre-miRNAs, the 3′ adapter was ligated and PCR was performed using a pre-miRNA specific forward primer and a reverse primer complementary to the 3′ adapter sequence.

[2] This experiment was performed by Inha Heo

# CHAPTER II

# Results

## II-1. Human Dicer recognizes the 5′-phosphorylated end

In an attempt to understand the impact of pre-miRNA uridylation on Dicer processing, I prepared synthetic pre-let-7a-1 with extra uridine residues at the 3′ end (**Figure II-1A**, left). The RNA was labeled at the 5′ end with [$\gamma$-$^{32}$P] ATP (**Figure II-1A**, left, asterisk) and incubated with immunopurified human Dicer. With the 3′ elongated substrates, I expected to observe a shift of the cleavage site, which would yield shorter products from the 5′ strand. Surprisingly, the size of the major cleavage products from the variants remained the same (22 nt) (**Figure II-1A**, purple arrowheads), indicating that the pre-let-7a-1 variants were cleaved at the same site regardless of the 3′ extension. I observed similar cleavage patterns when synthetic pre-miR-16-1 and its variants were used (**Figure II-2A**). To exclude the possibility that the 3′ extension was trimmed back by a contaminating nuclease, I used an additional set of substrates labeled at the 3′-most internucleotide phosphate (**Figure II-2B**). These substrates were also cleaved at the same sites, indicating that Dicer functions independently of the 3′ end in determining the cleavage site. This processing pattern was not influenced by the sequences of the nucleotides added to the 3′ overhang: addition of adenosine or cytidine

**Figure II-1. In vitro Dicer processing assay using uridylated substrates**

(A) Synthetic pre-let-7a-1 (Let-7) with extra uridines at the 3′ end were $^{32}$P-radiolabeled at the 5′ end (asterisk) and then incubated with immunopurified Flag-tagged human Dicer. Numbers on top of RNA substrates indicate the length of 3′ overhang. Purple and green arrowheads indicate 5′ counting and 3′ counting cleavage products, respectively (holds true for all the purple and green arrowheads throughout the paper). Reaction buffer contains 2 mM $Mg^{2+}$ throughout this work, unless indicated otherwise.

(B) Dicer processing of synthetic dsRNAs (ds-35) with extra uridines at the 3′ end. Gray arrowhead indicates the cleavage product from the opposite end, which is applicable throughout the paper.

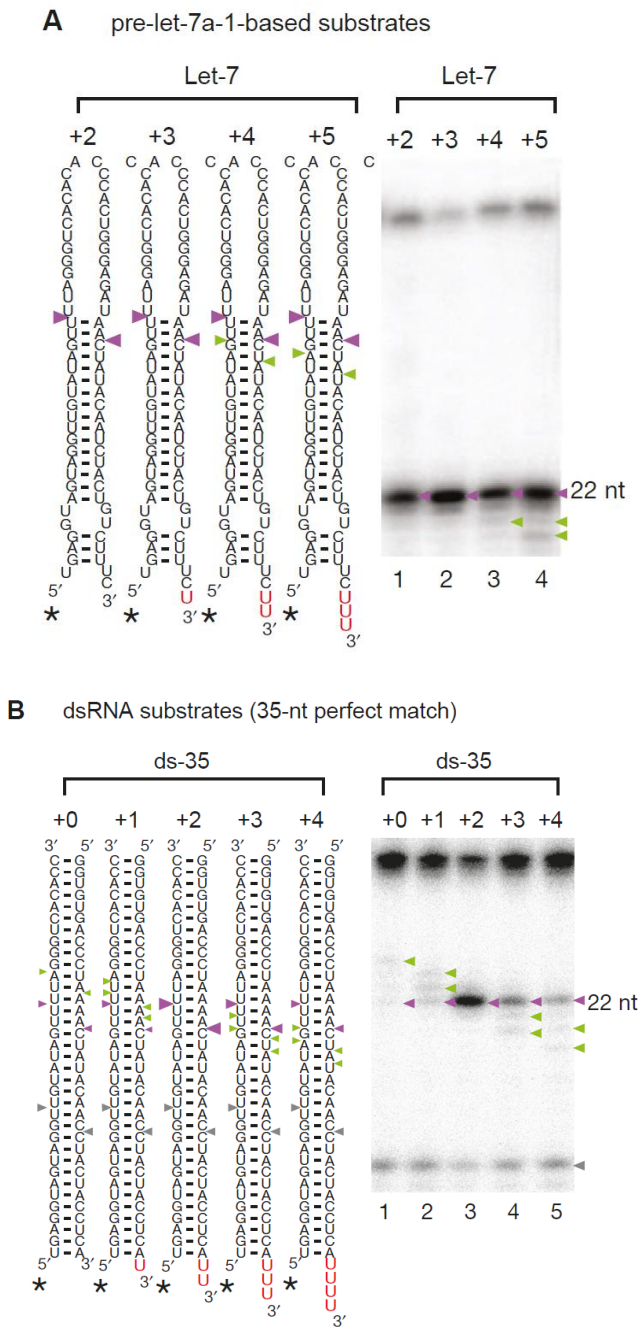
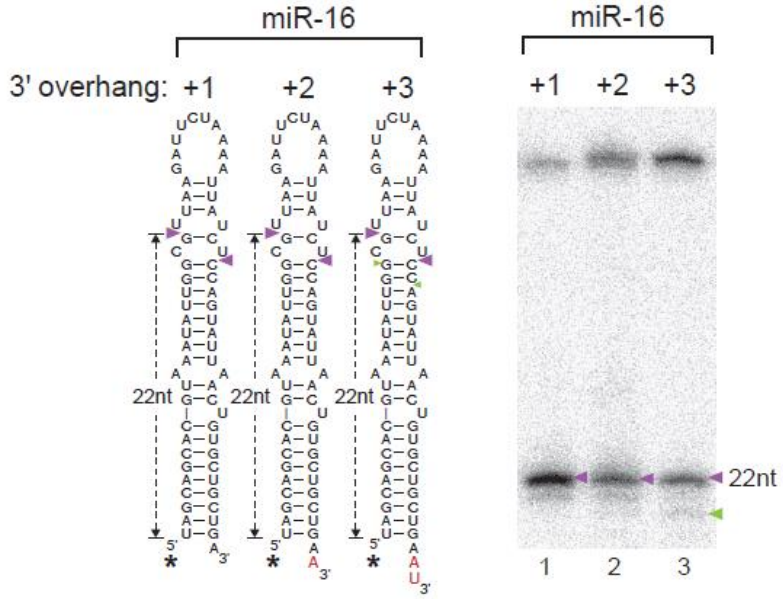

A
pre-let-7a-1-based substrates
Let-7
Let-7
+2
+3
+4
+5
+2
+3
+4
+5
22 nt
1
2
3
4
B
dsRNA substrates (35-nt perfect match)
ds-35
ds-35
+0
+1
+2
+3
+4
+0
+1
+2
+3
+4
22 nt
1
2
3
4
5

**Figure II-2. In vitro Dicer processing assay using different substrates**

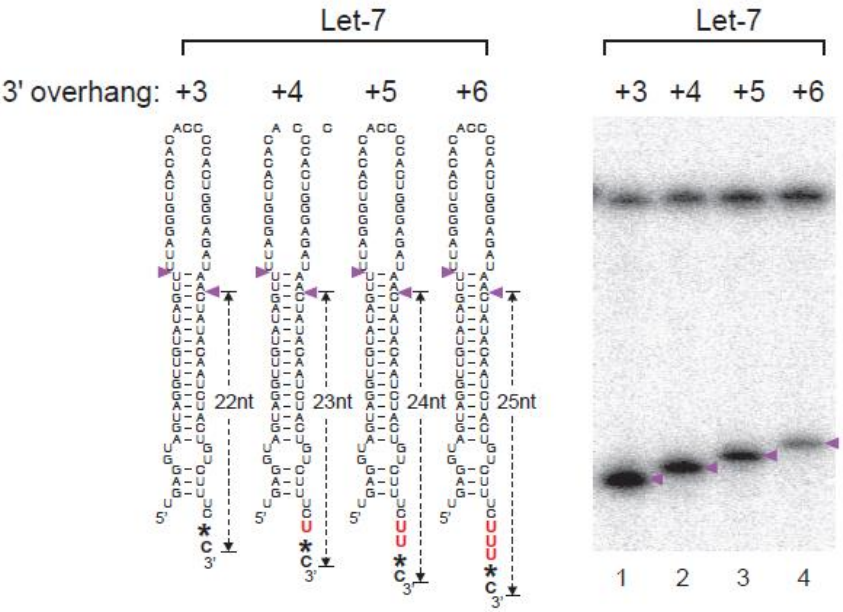
(A) Synthetic pre-miR-16-1 and its variants with an extended 3′ overhang were $^{32}$P-radiolabeled at the 5′ end (asterisk) and incubated with immunopurified Flag-tagged human Dicer

(B) The 3′ ends of pre-let-7a-1 substrates were labeled by [$\alpha^{32}$P]-pCp ligation (asterisk) and were incubated with immunopurified human Dicer. (B) Dicer processing of synthetic dsRNAs (ds-35) with extra uridines at the 3′ end. Gray arrowhead indicates the cleavage product from the opposite end, which is applicable throughout the paper.

**A** pre-miR-16-1 based substrates




**B** pre-let-7a-1 variants labeled at the 3' end



instead of uridine gave comparable results (**Figure II-2A** and data not shown).

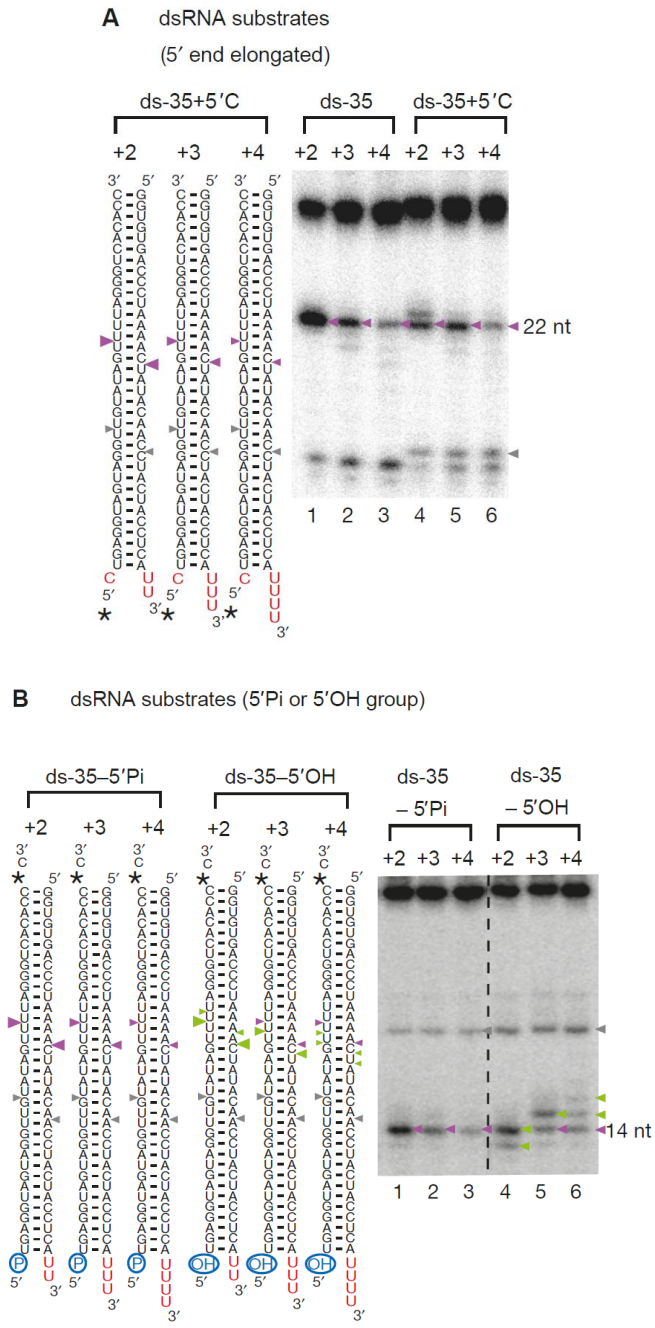
I next examined duplex RNAs with varying 3′ overhangs (**Figure II-1B**, left). One strand of the duplex was labeled with $^{32}$P at the 5′ end (**Figure II-1B**, left, asterisk). The predominant products from these dsRNAs were 22 nt in length in spite of the differences at the 3′ overhangs (**Figure II-1B**, purple arrowheads). Even the duplex with a 4 nt 3′ overhang was cleaved at the same site as the duplex with a 2 nt overhang, although another group of minor products was also seen (**Figure II-1B**, green arrowheads). These minor products were shortened as the 3′ overhang was elongated, indicating that the cleavage site moved towards the extending overhang. This cleavage pattern is expected of the 3′ counting model. I noticed that small amounts of 3′ counting products were also generated from pre-miRNAs (**Figure II-1A** and **Figure II-2A**, green arrowheads).

To summarize, I observed two types of Dicer cleavage events occurring in parallel. In one type, which is predominant for the substrates used here, the cleavage site does not change upon 3′ end elongation (marked with purple arrowheads). In another type, the cleavage site is determined based on the distance from the 3′ end (marked with green arrowheads). These results suggest that in addition to the 3′ end, Dicer may recognize additional structural feature(s) of the RNA substrate.

**Figure II-3. Dicer recognizes 5′ phosphorylated end for cleavage site selection**

(A) The 5′ end of ds-35 was extended by 1 nt (ds-35+5′C). Both the ds-35 and ds-35+5′C substrates were measured 22 nt from the 5′ end.

(B) The ds-35 substrates with either 5′ terminal phosphate or hydroxyl group (presented as cyan circle) were labeled at the 3′ end of the unvarying terminus by [α-$^{32}$P] pCp ligation (asterisk) and incubated with immunopurified human Dicer. Though the juxtaposed lanes are not contiguous, all of them are from a single gel, which is true for all the images with dashed lines throughout the paper


A
dsRNA substrates
(5′ end elongated)
ds-35+5′C
ds-35
ds-35+5′C
+2
+3
+4
+2 +3 +4 +2 +3 +4
22 nt
1 2 3 4 5 6
B
dsRNA substrates (5′Pi or 5′OH group)
ds-35–5′Pi
ds-35–5′OH
ds-35
– 5′Pi
ds-35
– 5′OH
+2 +3 +4 +2 +3 +4
14 nt
1 2 3 4 5 6

Of note, the substrates with a 2 nt 3′ overhang were cleaved most uniformly and efficiently, suggesting that Dicer binds to these canonical substrates most strongly by utilizing both 3′-dependent and 3′-independent mechanisms. When the terminal structure deviates from the optimal 2 nt 3′ overhang, either one of the two mechanisms seems to be used by Dicer, yielding the mixture of two distinct product populations.

I next questioned what determines the 3′-independent selection of the cleavage site. Because the cleavage site is always 22 nt away from the 5′ end (**Figure II-1**), I presumed that the 5′ end may play a role. To test this idea, dsRNA substrates were extended by 1 nt at the 5′ end (**Figure II-3A**, one additional cytidine at the 5′ end). Dicer still yielded products of ~22 nt from these substrates (**Figure II-3A**, purple arrowheads), indicating that the cleavage site shifted by 1 nt when the 5′ end was extended. Hence, Dicer measures a set distance (~22 nt) from the 5′ end. I refer to this as the '5′ counting model.'

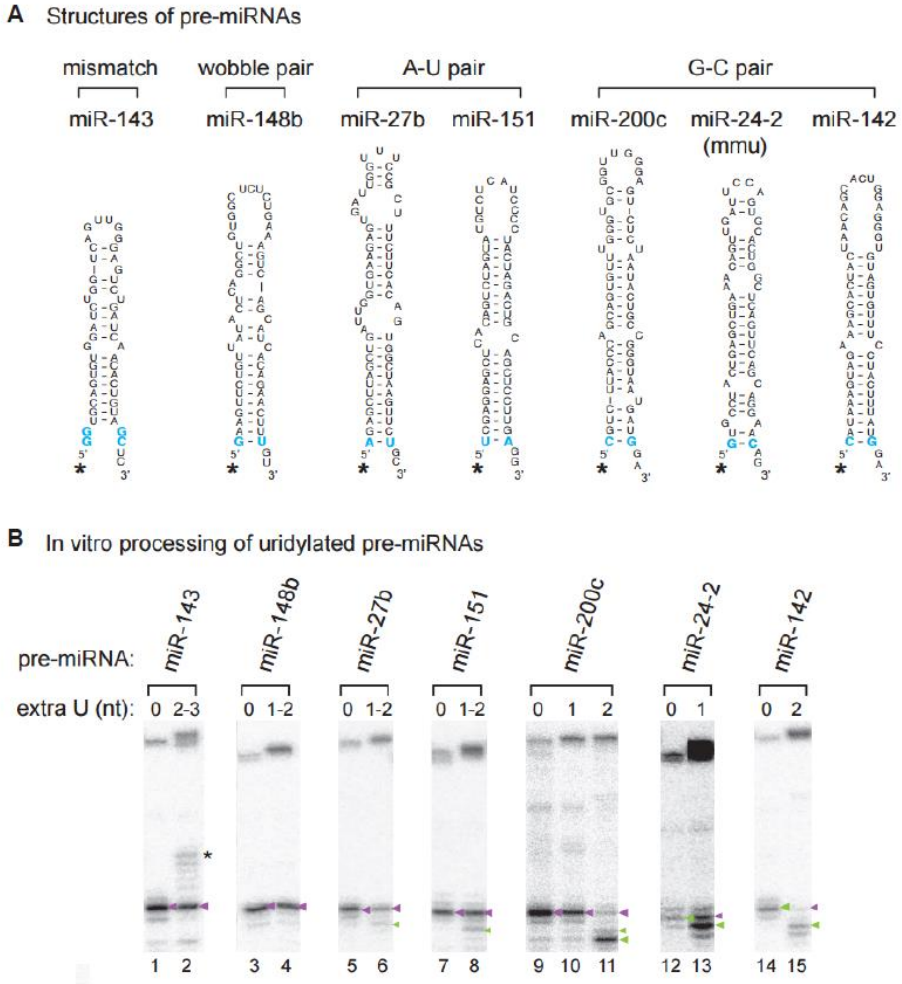
Because endogenous substrates of Dicer carry a 5′ terminal phosphate group, I tested whether the 5′ phosphate plays any role in the recognition of the 5′ end. For this, two sets of dsRNA substrates were generated. One set carries a 5′ terminal phosphate while the other set contains a hydroxyl group (**Figure II-3B**). Both sets were labeled at the 3′ end of the opposite terminus to detect the cleavage products (**Figure II-3B**, asterisk). The cleavage patterns

**Figure II-4. Dicer processing of various substrates with different terminal structures[3]**

(A) Structures of pre-miRNAs used in (B). Terminal base-pairs are colored in cyan.

(B) Unmodified and uridylated pre-miRNAs were incubated with Flag-tagged human Dicer immunoprecipitate. Uridylated pre-miR-200c substrates were chemically synthesized, while other uridylated substrates were generated by in vitro uridylation reaction. Asterisk marks non-specific degradation products degradation products.

[3] These experiments were performed in collaboration with Inha Heo


A Structures of pre-miRNAs
mismatch
wobble pair
A-U pair
G-C pair
miR-143
miR-148b
miR-27b
miR-151
miR-200c
miR-24-2 (mmu)
miR-142
B In vitro processing of uridylated pre-miRNAs
pre-miRNA:
miR-143
miR-148b
miR-27b
miR-151
miR-200c
miR-24-2
miR-142
extra U (nt):
0 2-3
0 1-2
0 1-2
0 1-2
0 1 2
0 1
0 2
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

of the two sets differed drastically. The phosphorylated dsRNAs followed the 5′ counting rule while the dsRNAs lacking the 5′ phosphate mainly obeyed the 3′ counting rule (**Figure II-3B**). I also noticed that the length of the products became more variable in the absence of the terminal phosphate. Taken together, our results imply that Dicer interacts with the 5′ terminal phosphate so as to precisely locate itself on the substrate.

To investigate whether the 5′ counting rule can be generalized for other pre-miRNAs, I performed the Dicer processing assay on additional pre-miRNAs with 1-3 extra uridine residues at the 3′ end (**Figure II-4**). Pre-miR-143, pre-miR-148b, pre-miR-27b and pre-miR-151 largely followed the 5′ counting rule, while pre-miR-200c showed mixed patterns (**Figure II-4B**). Pre-miR-24-2 and pre-miR-142 comply mainly with the 3′ counting rule (**Figure II-4B**). Hence, although the 5′ counting applies to most pre-miRNAs tested, the relative contribution of the 5′ and 3′ ends appears to vary among pre-miRNAs. I noticed that pre-miRNAs following the 3′ counting rule are relatively stable at the termini (**Figure II-4A**). Pre-miRNAs following the 5′ counting rule, on the other hand, have less stable structures at the termini (mismatch, G-U, or A-U pair) (**Figure II-4A**). Thus, Dicer may require a flexible (thermodynamically unstable) 5′ terminus to efficiently recognize the 5′ end. To further test this notion, I changed $Mg^{2+}$ concentrations in our processing assays from the previous 2 mM to either lower (0.5mM) or higher

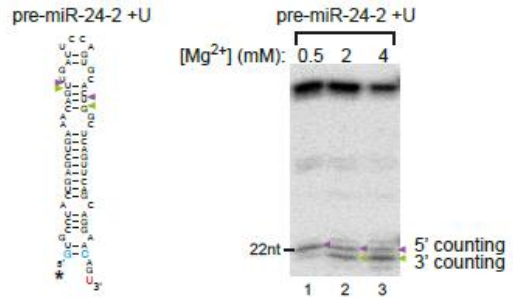
**Figure II-5. The 5′ counting rule prevails at low magnesium concentration**

(A) In vitro uridylated pre-miR-24-2 was incubated with the recombinant human Dicer proteins at different $Mg^{2+}$ concentrations (0.5 mM, 2 mM, and 4 mM). At low concentration of $Mg^{2+}$ (0.5 mM),the 5′ counting prevails.

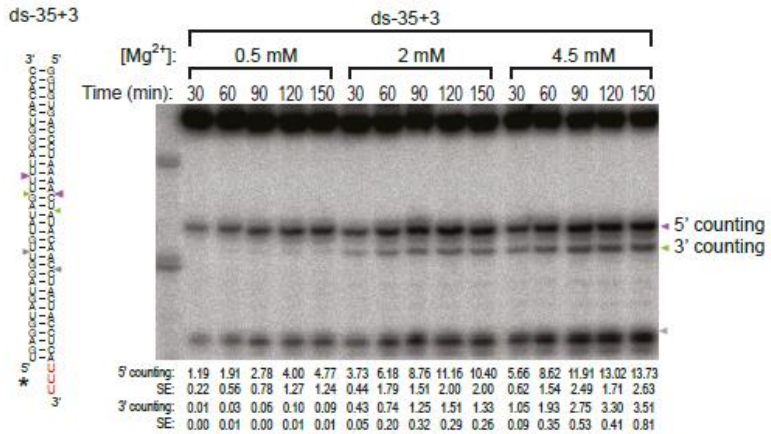
(B) The dsRNA with 3 nt overhang (ds-35+3) was incubated with immunoprecipitated Flag-Dicer at different magnesium ion concentrations (0.5 mM, 2 mM, 4.5 mM). The cleavage products were measured over time. Cleavage products from the 5′ counting and the 3′ counting are indicated as purple and green arrowheads, respectively. The value of product to substrate ratio and its standard error (SE) from three independent experiments are indicated below the gel image.

(C) Proportion of the 3′ counting product in b was measured. The average proportions of the 3′ products from triplicate experiments are shown next to the circle. Notably, much less 3′ counting product was generated at lower magnesium condition even when the amount of total cleavage product was comparable.

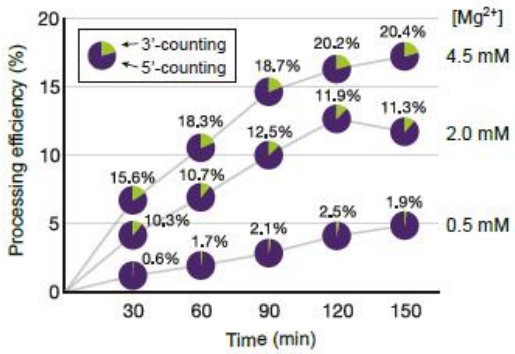
**A** Uridylated pre-miR-24-2 at different magnesium concentrations




**B** In vitro processing at different magnesium concentrations




**C** Proportion of 3' counting at different magnesium concentrations



(4 mM) concentrations, because $Mg^{2+}$ ion is known to stabilize dsRNA structure (Serra et al., 2002). $Mg^{2+}$ ion indeed had a significant influence on Dicer processing of pre-miR-24-2: the 3′ counting rule prevails at 4 mM while the 5′ counting rule predominates at 0.5 mM (**Figure II-5A**). A similar observation was made when a duplex RNA with a 3 nt overhang was used **Figure II-5B** and **Figure II-5C**). It is likely that at a low $Mg^{2+}$ concentration, the terminal stem region tends to unwind, thereby facilitating the 5′ end recognition by Dicer. Given that a physiological concentration of free $Mg^{2+}$ ion is estimated to be 0.5-1 mM (Gunther, 2006), the 5′ counting rule may apply to most 3′-modifed pre-miRNAs in vivo although I do not rule out the possibility that some pre-miRNAs with stable termini may follow the 3′ counting rule.

## II-2. Conservation of the 5′-counting rule

Since the 3′ counting model was proposed mainly from the work on Giardia Dicer1 (MacRae et al., 2007; Macrae et al., 2006), I examined Giardia Dicer using our substrates (**Figure II-6A**). Compared to human Dicer, Giardia Dicer yielded slightly larger products (24-26 nt) as previously observed (MacRae et al., 2007). Notably, while human Dicer generated mainly 5′ counting products with only minimal 3′ counting products (**Figure II-6A**, lanes 1-5), Giardia Dicer cleaved dsRNA substrates by strictly measuring from the 3′ end (**Figure II-6A**, lanes 6-10), which is consistent with the previous report (MacRae et al., 2007). Moreover, Giardia Dicer cleaved the blunt-ended substrate most efficiently while human Dicer showed a strong preference for the 2 nt 3′ overhang terminal structure. Thus, human Dicer differs significantly from Giardia Dicer in substrate recognition.

I also investigated the processing pattern of Drosophila Dicer-1. Dicer-1 acts in complex with the co-factor Loquacious-PB (Loqs-PB) (Forstemann et al., 2005; Jiang et al., 2005; Miyoshi et al., 2010; Saito et al., 2005). The Dicer-1/Loqs-PB complex is specialized for pre-miRNA processing while another Dicer (Dicer-2) and its cofactor R2D2 are responsible for siRNA processing (Lee et al., 2004). To prepare the Dicer-1 complex, Myc-tagged Loqs-PB was transiently expressed in Drosophila S2

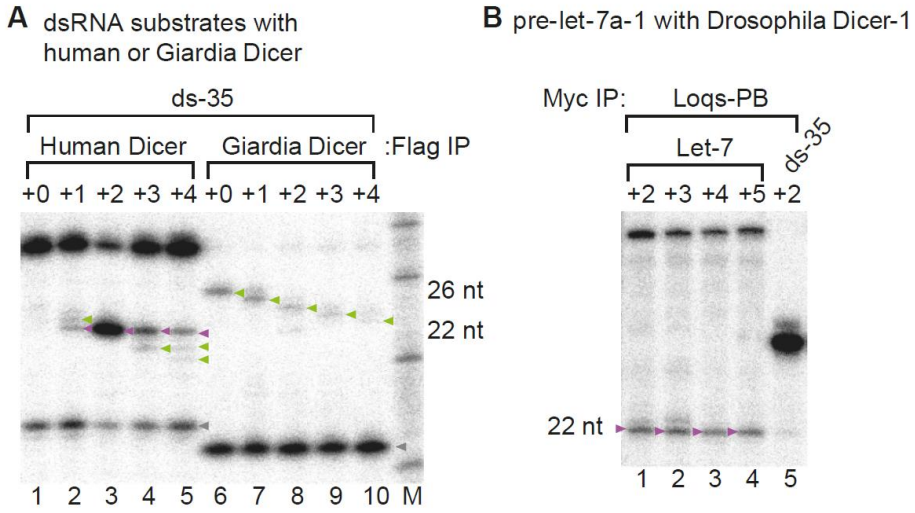



**Figure II-6. The 5′ counting rule is conserved in Drosophila Dicer-1 but not in Giardia Dicer**

(A) Flag-tagged human Dicer and Giardia Dicer were prepared by immunoprecipitation and incubated with the ds-35 substrates.

(B) Myc-tagged Drosophila Loquacious-PB (Loqs-PB) was expressed in Drosophila S2 cells followed by immunoprecipitation using anti-Myc antibody conjugated beads. The immunopurified Dicer-1/Loqs-PB complexes were incubated with Let-7 substrates. The ds-35 (+2) substrate was used as a negative control as fly Dicer-1 is specialized for pre-miRNA processing.

cells and immunoprecipitated using anti-Myc antibody. This complex was previously reported to cleave pre-miRNAs efficiently (Miyoshi et al., 2010; Saito et al., 2005). When pre-let-7a-1 variants were incubated with the Dicer-1/Loqs-PB complex, all variants were cleaved into 22 nt products, without any detectable products following the 3′ counting pattern (**Figure II-6B**). Taken together, the 5′ end recognition mechanism may be conserved in higher eukaryotes but not in organisms such as Giardia, which represents one of the earliest surviving branches of the eukaryotic phylogenetic tree.

## II-3. Identification of the 5′-recognition pocket

To identify the motif that binds to the 5′ end, I selected putative RNA interacting residues (basic, polar) located around the PAZ domain and mutated these residues to alanines (**Figure II-8A**). To narrow down the candidate residues, I predicted the 3D structure of the region encompassing the PAZ domain by applying I-TASSER simulation (Roy et al., 2010) (**Figure II-7**). Assuming that the 5′ end binding residues are located ~20 Å away from the 3′-pocket, which is the expected distance between the 5′ and 3′ ends of a 2 nt 3′ overhang structure, I selected putative RNA interacting residues found in this area (R811, R986, and R993). In addition, I also included stretches of

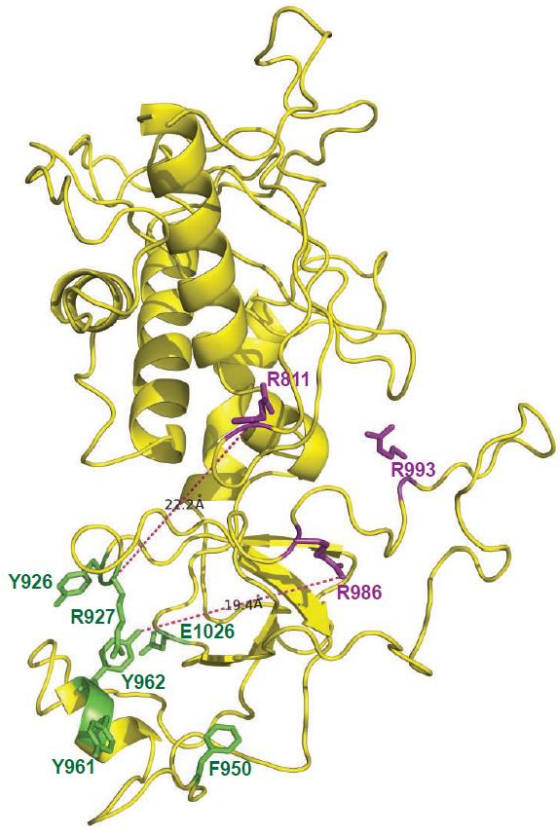


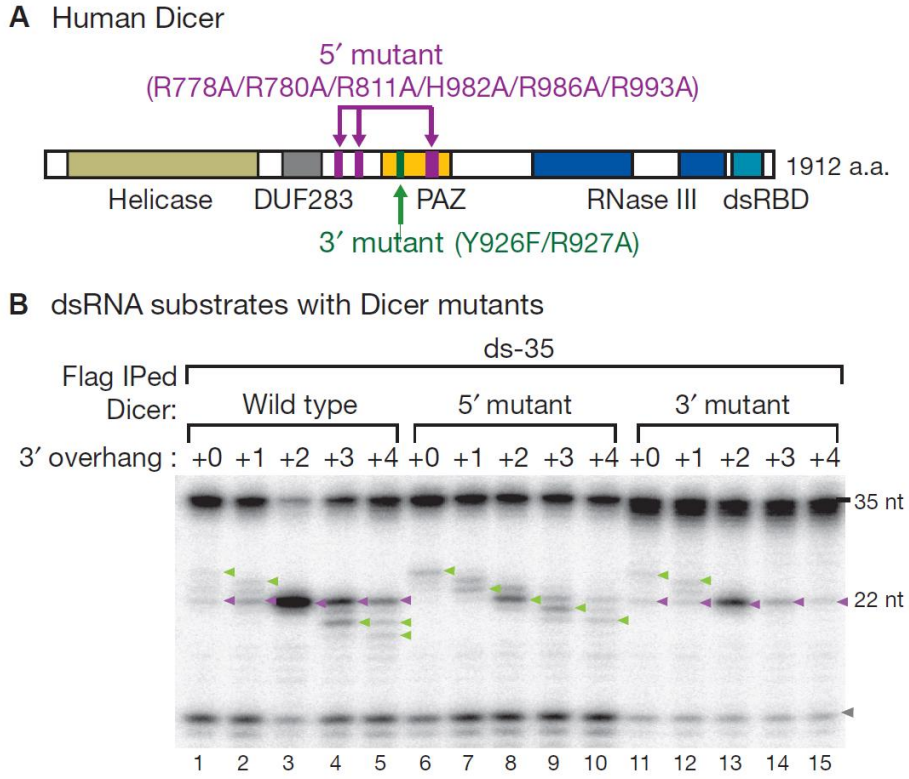

**Figure II-7. Three-dimensional model of Dicer fragment containing the PAZ domain**

The 3D structure of Dicer fragment (751-1070) was modeled by I-Tasser simulation. The conserved residues constituting the 3′-pocket of the PAZ domain (Y926, R927, F950, Y961, Y962 andE1026) are shown in green. Residues which bind to the 5′end of RNA are shown in purple. The predicted distances between the 3′-pocket and 5′end-binding residues are presented.

**Figure II-8. Mapping residues comprising the 5′-pocket**

(A) Domain organization of human Dicer. The mutated sites in the 5′-pocket and the 3′-pocket are shown as purple and green bars, respectively, and the mutations are listed.

(B) Wild-type and mutant Dicer proteins were immunopurified and incubated with ds-35 substrates. The 5′-pocket mutant is impaired in 5′ counting while the 3′-pocket mutant is defective in 3′ counting.


A Human Dicer
5′ mutant
(R778A/R780A/R811A/H982A/R986A/R993A)
1912 a.a.
Helicase
DUF283
PAZ
RNase III
dsRBD
3′ mutant (Y926F/R927A)
B dsRNA substrates with Dicer mutants
ds-35
Flag IPed Dicer:
Wild type
5′ mutant
3′ mutant
3′ overhang :
+0 +1 +2 +3 +4 +0 +1 +2 +3 +4 +0 +1 +2 +3 +4
35 nt
22 nt
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

basic or polar residues which are conserved between human and Drosophila Dicer-1.

As a result, the mutants at R778/R780, R811, and R986/R993 produced a significantly less amount of 5′ counting products, indicating that these mutants are defective in 5′ end recognition. When I combined these mutations to generate "5′-mutant" (R778A/R780A/R811A/H982A/R986A/R993A), the cleavage pattern clearly shifted to the 3′ counting one (**Figure II-8B**, lanes 6-10). The change in the cleavage pattern is highly specific to the identified residues; the point mutations at S984, H994, and W1014, which are located closely to the 5′-interacting residues, did not affect the cleavage pattern. Thus, our results indicate that a basic motif composed of R778, R780, R811, R986, and R993 (5′-pocket) is required for 5′ end recognition. These amino acids are conserved in Drosophila Dicer-1 but not in Giardia Dicer (**Figure II-9**).

As a control, I introduced mutations at Y926 and R927, which are conserved and located in the 3′-pocket of the PAZ domain (**Figure II-8A**). This mutant (3′-mutant) produced 22 nt products while losing most of the 3′ counting products, indicating that the 3′ counting mechanism is disrupted in this mutant (**Figure II-8B**, lanes 11-15). This result is consistent with previous findings that 3′ counting is dependent on the interaction between the 3′ end and the PAZ domain (MacRae et al., 2007; Zhang et al., 2004).

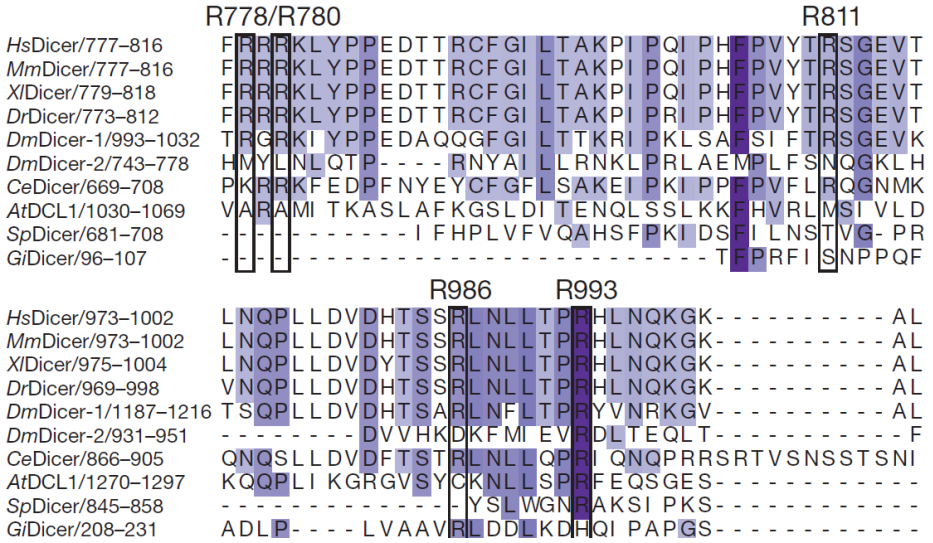



**Figure II-9. Conservation of 5′-pocket**

Amino acid sequences of the Dicer proteins from various species (Hs, Homo sapiens; Mm, Mus musculus; Xl, Xenopus laevis; Dm, Drosophila melanogaster; Ce, Caenorhabditis elegans; At, Arabidopsis thaliana; Sp, Schizosaccharomyces pombe; Gi, Giardia intestinalis) are aligned using ClustalX program and the region spanning the 5′-pocket is presented. The 5′ end binding residues are indicated with boxes.

Overall processing efficiency was reduced in the 5′-mutant as well as in the 3′-mutant (**Figure II-8B**). This supports the notion that human Dicer utilizes both the 5′ end and the 3′ end for substrate binding. It is noted that the substrate with a 2 nt overhang was cleaved more heterogeneously by the 5′-mutant Dicer (21-23 nt) compared to wild-type Dicer (22 nt) (**Figure II-8B**, compare lanes 3 and 8), indicating that the 5′-mutant lacks precision in processing. This result is consistent with our finding that the 5′ phosphate is required for production of uniform products (**Figure II-3B**). Altogether, our in vitro data suggest that 5′ end recognition is important not only for 3′-modified pre-miRNAs but also for canonical substrates such as unmodified pre-miRNAs.

## II-4. The 5′-pocket is required for miRNA biogenesis

To evaluate the biological relevance of my findings, I introduced Dicer expression plasmids (wild-type and 5′-mutant) into Dicer-null embryonic stem (ES) cells (Murchison et al., 2005). The small RNA populations from two biological replicates were sequenced by Illumina Genome Analyzer II (**Figure II-10** and **Table II-1**). The wild-type Dicer protein successfully replenished the miRNA pool while the 5′-mutant showed significant defects. When the 5′-mutant was expressed, the overall miRNA abundance was

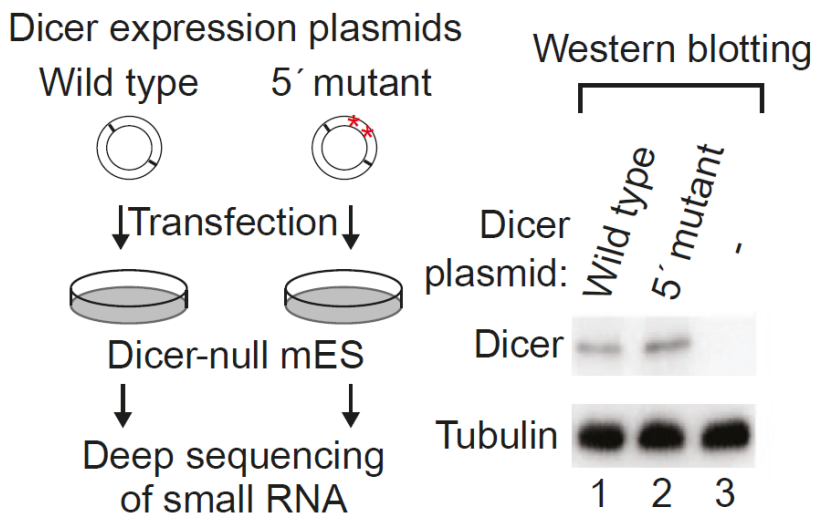



**Figure II-10. Experimental scheme for Dicer rescue assay**

(Left) Expression plasmid for either wild-type or 5′-mutant Dicer was transfected into Dicer-null mouse ES cells. Small RNAs were extracted from the cells and analyzed by deep-sequencing (two biological replicates for wild-type and the 5′-mutant). (Right) Western blotting of the first experimental set using anti-Dicer antibody shows that wild-type and mutant Dicer were expressed at comparable levels. The tubulin protein was detected as a loading control. The second set also expressed the Dicer proteins at comparable level (data not shown).

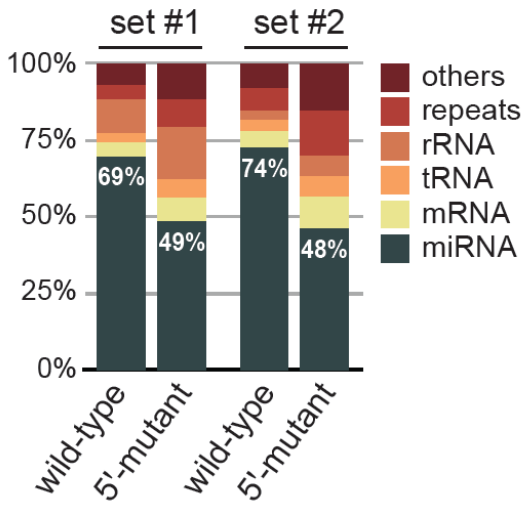
**Figure II-11. MiRNA levels are reduced in the 5′-mutant-rescued ES cells[4]**

(A) Genomic annotation of small RNA sequencing libraries from the rescued cells. The mutant libraries contained a significantly lower number of miRNA reads compared to the wild-type libraries. Note that reads from other RNAs were proportionately increased, compensating for the reduction of miRNAs in the library.

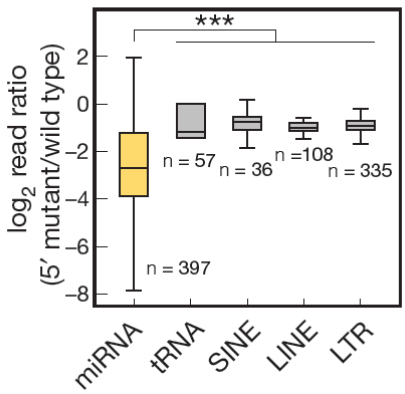
(B) Transcripts mapped to miRNA loci specifically decreased in the mutant library (P58.13310259, Mann–Whitney U-test, first replicate). Box represents the first and third quartiles and the internal bar indicates the median. Whiskers denote the lowest and highest values within 1.53interquartile range of the first and third quartiles, respectively. *n* represents the number of transcripts.

[4] Data analysis was performed by Hyeshik Chang

**A** Composition of small RNA libraries




**B** Reduction of miRNAs



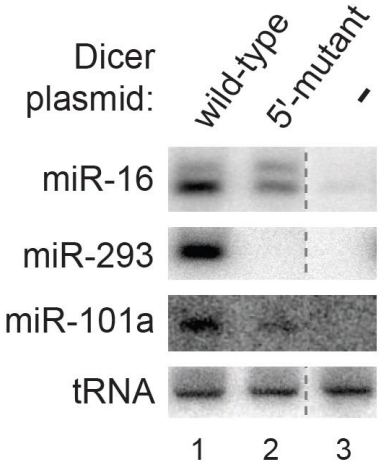



**Figure II-12. Northern blotting of miRNAs in the 5′-mutant-rescued ES cells**

Northern blotting confirms that the levels of mature miRNAs were lower in the 5′-mutant rescued cells than in the wild-type rescued cells.

significantly lower compared to values observed with wild-type Dicer (**Figure II-11**). Other RNA species such as tRNAs remained unaffected, indicating the differences are specific to miRNAs (**Figure II-11**). The drastic reduction of miRNA abundance was further confirmed by northern blotting of miR-16, miR-293, and miR-101a (**Figure II-12**).

I next examined the impact of the 5′-pocket mutation on processing site selection. As the 3′ ends of small RNAs are known to be frequently modified after the Dicer processing step (Burroughs et al., 2010; Chiang et al., 2010; Heo et al., 2008; Wu et al., 2009), I used the 5′ end of miRNAs (or miRNAs*) to infer cleavage sites. Drosha creates the 5′ end of 5′ strand miRNAs (5p miRNAs) while Dicer makes the 5′ end of 3′ strand miRNAs (3p miRNAs) (**Figure II-14**, left). When the cleavage sites from wild-type and 5′-mutant libraries were compared, ~35 % of miRNAs showed significant changes in Dicer cleavage sites (41 out of 117; below 5 % false discovery rate) (**Figure II-14**, right panel, **Figure II-13**). On the contrary, Drosha processing sites remained largely unchanged (**Figure II-14**, left panel and **Figure II-15**), indicating that the differences in the small RNA population are due to the mutation in Dicer. The changes in Dicer cleavage sites often led to seed alterations and/or strand switches (**Table II-2**). The results from deep sequencing are highly consistent with the results of the in vitro assays.

To further confirm our findings, I carried out in vitro processing of pre-miR-30a and pre-miR-200c, which showed significant changes in the ES

cell rescue experiments (**Figure II-16**). The 5′-mutant Dicer was markedly impaired in both efficiency and accuracy of pre-miRNA processing in vitro (**Figure II-16**). The 3′-pocket mutation reduced processing activity without significantly altering cleavage site selectivity (**Figure II-16**).

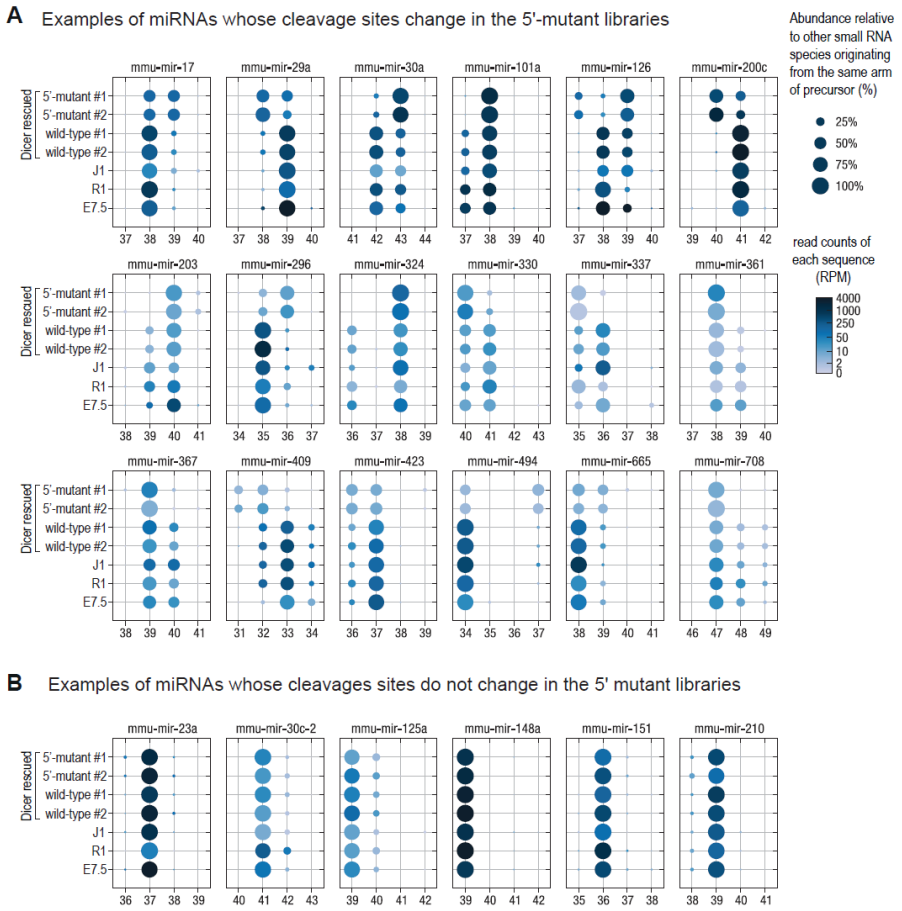
To compare the function of 5′- and 3′-pockets in vivo, I also expressed 3′-pocket mutant in Dicer knockout ES cells, performed small RNA sequencing, and compared with the result from 5′-pocket mutant rescue experiments. Interestingly, miRNAs which are highly affected in 5′-pocket mutant also tend to be affected in 3′-pocket mutant (r=0.55, **Figure II-17A**), suggesting that these miRNAs require both pockets for efficient processing. Dicer cleavage sites were more dramatically changed in 5′-pocket mutant than 3′-pocket mutant (**Figure II-17B**), consistent with in vitro data which showed that 5′-pocket is more important for accurate cleavage site selection. Taken together, the 5′-pocket is critical for efficient and precise generation of miRNA.

**Figure II-13. Dicer cleavage patterns in the Dicer-rescued libraries and the control libraries**[5]

(A) Comparison of the Dicer cleavage sites in different libraries. The Dicer cleavage sites from the Dicer-rescued cells are presented along with J1 ES cell line, R1 ES cell line, and mouse embryo at day 7.5. The x-axis represents the position of Dicer cleavage site on the pre-miRNA hairpin counted from the 5′ end. The dot represents the 3p miRNA whose 5′ end starts at the marked position. The size of the circle indicates the percentage of the miRNA isoforms whose 5′ end matches the given position. The color of the circle indicates the reads per million (RPM) of the miRNA species.

(B) Some miRNAs do not show significant changes of cleavage sites in the 5′-mutant libraries compared to the wild-type libraries.

[5] Data analysis was performed by Hyeshik Chang


A
Examples of miRNAs whose cleavage sites change in the 5'-mutant libraries
mmu-mir-17
mmu-mir-29a
mmu-mir-30a
mmu-mir-101a
mmu-mir-126
mmu-mir-200c
mmu-mir-203
mmu-mir-296
mmu-mir-324
mmu-mir-330
mmu-mir-337
mmu-mir-361
mmu-mir-367
mmu-mir-409
mmu-mir-423
mmu-mir-494
mmu-mir-665
mmu-mir-708
Dicer rescued
5'-mutant #1
5'-mutant #2
wild-type #1
wild-type #2
J1
R1
E7.5
Abundance relative to other small RNA species originating from the same arm of precursor (%)
25%
50%
75%
100%
read counts of each sequence (RPM)
4000
1000
250
50
10
2
0
B
Examples of miRNAs whose cleavages sites do not change in the 5' mutant libraries
mmu-mir-23a
mmu-mir-30c-2
mmu-mir-125a
mmu-mir-148a
mmu-mir-151
mmu-mir-210

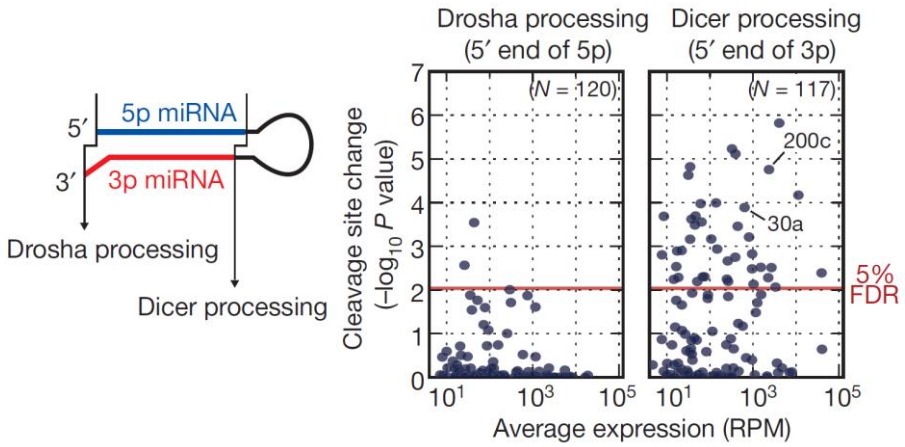



**Figure II-14. The 5′-pocket is critical for accurate cleavage in vivo**[6]

Left, Drosha and Dicer cleavage sites were inferred from the 5′ end of 5p and 3p miRNAs, respectively. Right, the dissimilarity of cleavage pattern was quantified by Kullback–Leibler divergence (KLD), and statistical significance was measured using two-sample t-test. Red line corresponds to 5% false discovery rate (FDR). RPM, reads per million.

[6] Data analysis was performed by Hyeshik Chang

**Figure II-15. Dicer cleavage site of miRNAs is shifted only in the 5′-mutant libraries**[7]

(A) Dicer and Drosha cleavage sites from the wild-type or 5′-mutant libraries were compared with those from the control libraries (deep sequencing from J1 and R1 ES cell lines that express normal level of Dicer). The difference of the cleavage site between the libraries was calculated by Kullback-Leibler divergence (KLD). The x-axis represents the score of cleavage site change quantified by KLD between the wild-type and the control libraries. The y-axis represents the score of cleavage site change between the mutant and the control libraries. The red and blue dots represent Dicer and Drosha cleavage site change of each miRNA, respectively. The red dots are shifted to upper-left region, indicating that Dicer cleavage site is altered specifically in the mutant libraries.

(B) Comparison of the cleavage site change between libraries from two biological replicates. It shows that the replicates are highly similar to each other and that the results are reproducible. The cleavage site change between each replicate and R1 library was used for comparison.

---

[7] Data analysis was performed by Hyeshik Chang

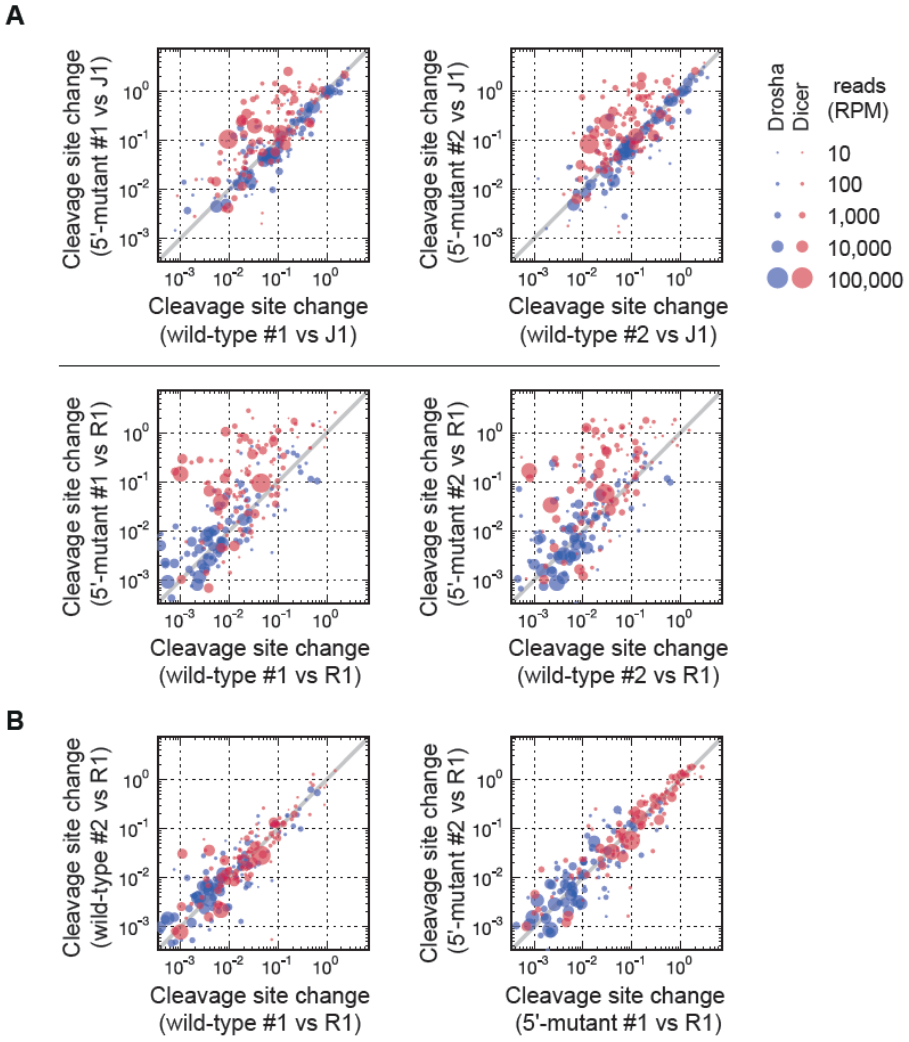
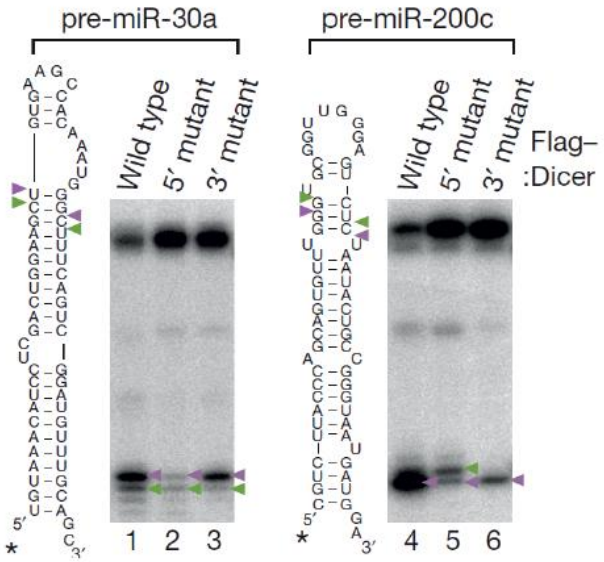

A
Cleavage site change (5'-mutant #1 vs J1)
Cleavage site change (wild-type #1 vs J1)
Cleavage site change (5'-mutant #2 vs J1)
Cleavage site change (wild-type #2 vs J1)
Drosha
Dicer
reads (RPM)
10
100
1,000
10,000
100,000
Cleavage site change (5'-mutant #1 vs R1)
Cleavage site change (wild-type #1 vs R1)
Cleavage site change (5'-mutant #2 vs R1)
Cleavage site change (wild-type #2 vs R1)
B
Cleavage site change (wild-type #2 vs R1)
Cleavage site change (wild-type #1 vs R1)
Cleavage site change (5'-mutant #2 vs R1)
Cleavage site change (5'-mutant #1 vs R1)

**Figure II-16. Validation of Dicer cleavage site shift by in vitro processing**

(A) Pre-miR-30a and pre-miR-200c were incubated with the same amount of wild-type or mutant Dicer protein.

(B) The hairpin sequences of pri-miR-30a and pri-miR-200c are shown. The 5p and 3p miRNA sequences are marked in blue and red, respectively, according to miRbase annotation. The read numbers (RPM) of two most frequent 3p sequences from each library are presented on the right side. The ratio between two most frequent isoforms is indicated below the read numbers.

**A** *In vitro* processing of pre-miRNAs




**B** sequencing result

pre-miR-30a

| | reads (RPM) | | | |
|---|---|---|---|---|
| | wild-type | | 5'-mutant | |
| | #1 | #2 | #1 | #2 |
| 5p 3p | | | | |
| GCGACUGUAAACAUCCUCGACUGGAAGCUGUGAAGCCACAAAUGGGCUUUCAGUCGGAUGUUUGCAGCUGC | | | | |
| (((.(((((((((((..(((((((((((((( ....))).....)))))))))))))))))))))).))) | | | | |
| .....CUUUCAGUCGGAUGUUUGCAGC... | 289 | 330 | 34 | 118 |
| .....UUUCAGUCGGAUGUUUGCAGC... | 125 | 132 | 168 | 583 |
| ratio between two isoforms: | 2.31 | 2.50 | 0.20 | 0.20 |

pre-miR-200c

| | | | | |
|---|---|---|---|---|
| 5p 3p | | | | |
| CCCUCGUCUUACCCAGCAGUGUUUGGGUGCUGGUUGGGAGUCUCUAAUACUGCCGGGUAAUGAUGGAGG | | | | |
| ((..((((((((((.(((((((((.(((.((( .....)))))).)))))))).))))))).)))))..)) | | | | |
| .....CUAAUACUGCCGGGUAAUGAUGGA.. | 13 | 40 | 276 | 1049 |
| .....UAAUACUGCCGGGUAAUGAUGGA.. | 1889 | 4578 | 168 | 593 |
| ratio between two isoforms: | 0.01 | 0.01 | 1.64 | 1.77 |

**Figure II-17. Comparison of 5′-pocket mutant and 3′-pocket mutant**

(A) MiRNA expression fold change in each pocket mutant libraries

(B) Dicer cleavage site heterogeneity is measured by calculating the difference between Dicer cleavage patterns of mutant and control libraries. As a control for natural variation in Dicer cleavage patterns, I calculated Dicer cleavage site heterogeneity between WT samples.

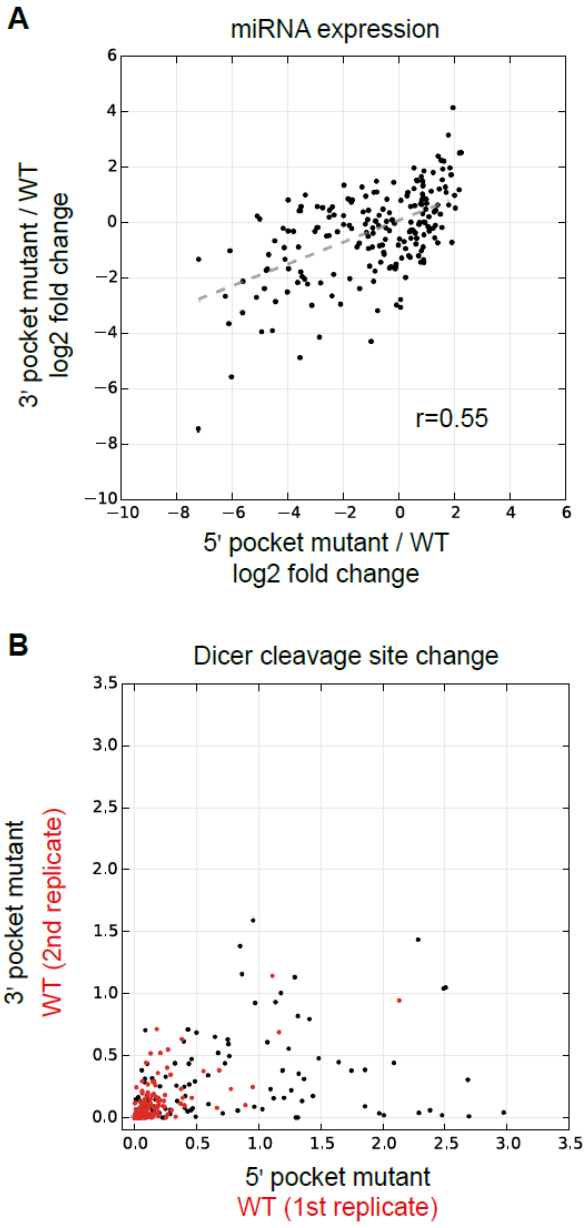

A
miRNA expression
3' pocket mutant / WT
log2 fold change
r=0.55
5' pocket mutant / WT
log2 fold change
B
Dicer cleavage site change
3' pocket mutant
WT (2nd replicate)
5' pocket mutant
WT (1st replicate)

**Table II-1. Sequencing read statistics**

Numbers of sequence reads in steps of data processing and basic information for each library. The “original total reads” row corresponds to the total read counts provided in FASTQ format by Illumina® Genome Analyzer. Then low quality reads were removed, remaining read counts are indicated in “- low quality reads”. Low complexity reads and too short reads were filtered (“- low complexity / short reads”), adapter and primer sequences were depleted from the library (“-technical contaminants”). BWA mapped the remaining reads to mm9 genome assembly (“genome-mappable reads”) with maximum four mismatches. The genome coordinates for reads were intersected with miRBase 17 and UCSC versions of RefSeq and RepeatMasker annotations as of Apr 8th 2011 (“mapped transcripts”). Descriptive statistics for read length distribution in five rows at the bottom were calculated from the low quality-filtered libraries.

| | AB2.2 Dicer wild-type biological replicate #1 | AB2.2 Dicer 5' mutant biological replicate #1 | AB2.2 Dicer wild-type biological replicate #2 | AB2.2 Dicer 5' mutant biological replicate #2 | R1 control | J1 control (Babiarz et al.) | mEB e7p5 (Chiang et al.) |
|---|---|---|---|---|---|---|---|
| **GEO Accession** | GSM689053 | GSM689055 | GSM689054 | GSM689056 | GSM689057 | GSM314552 | GSM510465-8 |
| **Original total reads** | 20,927,677 | 26,388,569 | 20,542,593 | 18,045,394 | 23,711,709 | N/A | N/A |
| **- low quality reads** | 19,347,023 | 26,127,423 | 20,316,835 | 17,228,160 | 19,364,022 | 7,492,936 | 3,942,262 |
| **- low complexity / short reads** | 16,595,277 | 19,111,851 | 19,659,009 | 14,925,101 | 15,595,416 | 5,977,898 | 3,588,747 |
| **- technical contaminants** | 16,543,477 | 18,974,922 | 19,603,398 | 14,799,379 | 15,575,233 | 5,974,234 | 3,560,566 |
| **Genome-mappable reads** | 16,222,487 | 18,331,076 | 19,556,185 | 14,743,506 | 15,427,206 | 5,955,416 | 3,522,038 |
| **Mapped transcript: miRNA** | 11,450,444 | 9,235,807 | 14,470,695 | 7,130,999 | 8,980,195 | 2,965,005 | 2,498,677 |
| **Mapped transcript: rRNA** | 1,780,346 | 3,043,456 | 596,272 | 954,016 | 1,094,628 | 386,832 | 142,595 |
| **Mapped transcript: tRNA** | 491,565 | 1,053,061 | 694,921 | 977,657 | 229,308 | 163,451 | 292,566 |
| **Mapped transcript: mRNA** | 237,731 | 431,769 | 340,225 | 540,102 | 421,135 | 45,286 | 34,718 |
| **Mapped transcript: intron** | 488,184 | 993,288 | 676,878 | 957,657 | 2,024,420 | 276,536 | 85,015 |
| **Mapped transcript: SINE** | 178,285 | 404,499 | 255,067 | 365,284 | 365,623 | 66,970 | 34,058 |
| **Mapped transcript: LINE** | 257,955 | 471,646 | 562,504 | 929,365 | 508,302 | 207,738 | 58,076 |
| **Mapped transcript: LTR** | 109,734 | 208,043 | 201,411 | 276,193 | 179,313 | 53,038 | 25,265 |
| **Read length average** | 20.22 | 18.52 | 21.99 | 21.21 | 20.84 | 19.96 | 21.41 |
| **Read length SD** | 5.49 | 7.11 | 1.83 | 2.64 | 4.10 | 2.74 | 2.46 |
| **Read length 1st quartile** | 20 | 17 | 22 | 19 | 19 | 18 | 20 |
| **Read length 2nd quartile** | 22 | 22 | 22 | 22 | 22 | 20 | 22 |
| **Read length 3rd quartile** | 23 | 23 | 23 | 23 | 23 | 22 | 23 |

**Table II-2. List of mature miRNAs with tendency of seed sequence changes by 5′-pocket site mutation**

Each row represents an annotated miRNA that has a unique sequence among all known murine miRNAs (miRBase 16). The most frequent seed sequence read ("mostfreq" in short) is determined from our wild-type small RNA sequencing libraries, and the second frequent seed sequence read ("2ndfreq" in short) is derived from the second abundant 5′ end of reads mapped to the miRNA. The changes of ratio between mostfreq and 2ndfreq in the 5′-mutant library indicate contribution of 5′-pocket site to processing of miRNA to produce correct seed sequence and target intentional targets.

| mature ID | arm | shift nt 2ndfreq | log2 mostfreq /2ndfreq change (mutant) | log mostfreq /2ndfreq change (control) | cleavage chg p (KLD) | log 5p/3p chg. (arm-switching) | mostfreq/ 2ndfreq wt #1 | mostfreq/ 2ndfreq wt #2 | mostfreq/ 2ndfreq mut #1 | mostfreq/ 2ndfreq mut #2 | mostfreq/ 2ndfreq ctl #1 | mostfreq/ 2ndfreq ctl #2 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| mmu-miR-142-3p | 3' | 1 | 7.23 | 0.09 | 1.51E-06 | 0.09 | 2.2 | 1.6 | 460.8 | 148.4 | 0.8 | 4.9 |
| mmu-miR-17* | 3' | 1 | -2.94 | 0.69 | 5.89E-06 | -0.06 | 7.6 | 7.1 | 1.0 | 1.0 | 5.4 | 23.7 |
| mmu-miR-29a | 3' | -1 | -4.10 | 1.58 | 7.68E-06 | 4.15 | 9.9 | 7.8 | 0.8 | 0.4 | 26.7 | 20.3 |
| mmu-miR-200c | 3' | -1 | -7.49 | 1.09 | 1.76E-05 | 2.93 | 84.6 | 72.4 | 0.5 | 0.4 | 68.8 | 320.9 |
| mmu-miR-1983 | 3' | 1 | -1.54 | -0.70 | 2.36E-05 | 7.77 | 6.5 | 4.7 | 1.5 | 1.2 | 3.3 | 2.6 |
| mmu-miR-293 | 3' | -1 | -4.61 | -0.41 | 6.76E-05 | 2.23 | 334.5 | 269.9 | 10.8 | 14.0 | 364.3 | 138.2 |
| mmu-miR-296-3p | 3' | 1 | -5.36 | -1.41 | 0.000102 | 6.30 | 14.2 | 16.9 | 0.4 | 0.5 | 7.6 | 4.3 |
| mmu-miR-28* | 3' | 1 | 3.64 | 0.20 | 0.000106 | -0.32 | 7.1 | 6.2 | 24.3 | 45.8 | 11.2 | 5.7 |
| mmu-miR-30a* | 3' | 1 | -3.92 | -0.58 | 0.00013 | 0.05 | 1.9 | 2.1 | 0.1 | 0.2 | 1.3 | 1.3 |
| mmu-miR-708* | 3' | 1 | 3.60 | -0.37 | 0.000207 | -0.31 | 2.5 | 2.9 | 7.3 | 4.6 | 3.4 | 1.8 |
| mmu-miR-423-3p | 3' | -1 | -1.97 | 0.26 | 0.00024 | 2.67 | 4.3 | 3.5 | 1.0 | 1.1 | 6.0 | 3.8 |
| mmu-miR-337-5p | 5' | -2 | -0.38 | -0.66 | 0.000287 | 1.06 | 18.2 | 13.8 | 3.0 | 5.9 | 28.2 | 3.3 |
| mmu-miR-494 | 3' | 3 | -4.95 | -0.03 | 0.000326 | 3.84 | 27.8 | 24.1 | 0.8 | 1.0 | 14.9 | 37.7 |
| mmu-miR-302b | 3' | 1 | 3.51 | 1.18 | 0.000349 | -0.82 | 2.9 | 2.8 | 26.6 | 33.6 | 4.9 | 8.3 |
| mmu-miR-126-3p | 3' | 1 | -3.62 | 0.83 | 6.15E-04 | 5.10 | 1.5 | 1.6 | 0.1 | 0.1 | 0.9 | 7.7 |
| mmu-miR-409-3p | 3' | -1 | -3.46 | -0.04 | 0.000682 | 3.27 | 2.5 | 2.8 | 0.3 | 0.3 | 2.8 | 2.3 |
| mmu-miR-25 | 3' | 1 | -3.56 | -0.57 | 0.001158 | 1.93 | 169.8 | 375.3 | 6.9 | 7.0 | 109.9 | 245.9 |
| mmu-miR-101a | 3' | -1 | 5.14 | -0.81 | 0.001499 | -0.49 | 4.3 | 3.5 | 95.7 | 151.1 | 3.0 | 1.6 |
| mmu-miR-361* | 3' | 1 | 5.46 | -1.77 | 0.001574 | -5.21 | 2.4 | 2.4 | 40.5 | 8.3 | 1.5 | 1.1 |
| mmu-miR-369-3p | 3' | -1 | 4.40 | 2.56 | 0.00178 | -1.07 | 7.1 | 5.2 | 91.6 | 39.3 | 18.9 | 45.3 |
| mmu-miR-410 | 3' | 1 | -4.43 | -1.22 | 0.002166 | 2.19 | 345.3 | 375.6 | 14.7 | 14.8 | 77.0 | 260.9 |
| mmu-miR-185 | 5' | 2 | -3.92 | -0.51 | 0.002723 | -2.22 | 146.0 | 139.5 | 5.4 | 5.4 | 74.9 | 205.6 |
| mmu-miR-127 | 3' | -1 | 3.65 | 0.41 | 0.003038 | 0.50 | 56.7 | 68.6 | 549.6 | 315.4 | 82.1 | 83.8 |
| mmu-miR-378 | 3' | 1 | -2.21 | 1.43 | 3.08E-03 | 3.46 | 10.6 | 12.7 | 1.7 | 3.8 | 25.1 | 36.4 |
| mmu-miR-381 | 3' | -1 | -5.80 | 0.58 | 0.00329 | 3.03 | 32.7 | 30.1 | 0.7 | 0.5 | 26.2 | 74.8 |
| mmu-miR-295 | 3' | 1 | 2.54 | 0.06 | 0.004078 | 0.65 | 4.2 | 4.9 | 24.4 | 28.5 | 4.9 | 4.6 |
| mmu-miR-20a* | 3' | 1 | -3.34 | 0.44 | 0.005005 | 0.88 | 5.9 | 1.5 | 0.5 | 0.2 | 4.5 | 2.8 |
| mmu-miR-135b* | 3' | 1 | -1.86 | 1.44 | 0.005006 | -0.06 | 2.3 | 1.4 | 0.6 | 0.5 | 6.0 | 3.3 |
| mmu-miR-19a | 3' | -1 | 2.79 | -0.54 | 0.005241 | -0.48 | 23.1 | 13.8 | 161.7 | 91.5 | 13.8 | 10.4 |
| mmu-miR-300 | 3' | 2 | 2.62 | 0.99 | 0.005653 | -2.06 | 16.3 | 12.0 | 37.4 | 29.1 | 21.9 | 33.4 |
| mmu-miR-503* | 3' | 1 | 3.30 | 0.63 | 0.005721 | -0.42 | 2.1 | 1.2 | 4.7 | 9.7 | 1.6 | 1.8 |
| mmu-miR-302d | 3' | 1 | 5.71 | -0.05 | 0.006387 | -1.35 | 16.3 | 15.1 | 564.1 | 148.6 | 18.2 | 13.9 |
| mmu-miR-541* | 3' | 2 | 4.45 | -0.14 | 0.006417 | -1.38 | 1.2 | 1.1 | 16.0 | 14.2 | 2.0 | 0.6 |
| mmu-miR-335-3p | 3' | -1 | -2.45 | 0.48 | 0.007428 | 1.43 | 10.0 | 4.8 | 1.6 | 1.0 | 8.7 | 10.0 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **mmu-miR-292-3p** | 3' | 1 | 1.64 | -0.42 | 0.008596 | -0.33 | 1.2 | 2.4 | 4.2 | 6.6 | 1.0 | 1.5 |
| **mmu-miR-340-5p** | 5' | -1 | -2.98 | -1.57 | 0.009872 | 2.58 | 207.9 | 77.1 | 12.3 | 12.2 | 40.2 | 36.5 |
| **mmu-miR-496** | 3' | -2 | -3.28 | -1.63 | 0.01179 | 4.42 | 5.0 | 6.6 | 0.5 | 1.0 | 1.7 | 2.1 |
| **mmu-miR-320** | 3' | 1 | 0.55 | -0.43 | 0.012757 | 0.17 | 22.8 | 23.5 | 27.1 | 40.2 | 12.3 | 23.4 |
| **mmu-miR-30b*** | 3' | 1 | -5.20 | 0.39 | 0.013084 | -0.10 | 16.5 | 19.1 | 0.6 | 0.7 | 3.6 | 32.1 |
| **mmu-miR-409-5p** | 5' | 11 | -4.11 | -1.12 | 0.013222 | 3.27 | 88.4 | 143.9 | 8.1 | 10.1 | 98.1 | 33.5 |
| **mmu-miR-335-5p** | 5' | 2 | -1.43 | 0.60 | 0.013505 | 1.43 | 13.6 | 12.4 | 7.4 | 3.0 | 22.5 | 12.6 |
| **mmu-miR-299** | 3' | 1 | -7.45 | -1.56 | 0.014096 | 7.60 | 1.1 | 2.4 | 0.0 | 0.0 | 0.2 | 1.7 |
| **mmu-miR-342-3p** | 3' | 2 | 1.25 | -0.06 | 0.015644 | -0.26 | 11.2 | 13.0 | 21.6 | 20.7 | 15.9 | 6.1 |
| **mmu-miR-872*** | 3' | -1 | 0.10 | -0.53 | 0.017295 | -0.88 | 33.1 | 30.4 | 9.5 | 4.8 | 47.9 | 15.7 |
| **mmu-miR-434-5p** | 5' | -1 | 1.05 | -0.03 | 0.019428 | 1.01 | 1.9 | 2.2 | 4.5 | 4.0 | 2.0 | 2.1 |
| **mmu-miR-101b** | 3' | -1 | 1.82 | -1.97 | 0.01963 | 1.20 | 21.8 | 16.6 | 48.8 | 70.3 | 3.3 | 7.3 |
| **mmu-miR-290-5p** | 5' | 1 | -2.87 | 0.43 | 2.46E-02 | -1.60 | 44.1 | 37.7 | 6.5 | 5.0 | 73.3 | 25.9 |
| **mmu-let-7d** | 5' | -1 | -1.00 | -0.83 | 0.025381 | -3.24 | 66.8 | 52.2 | 5.1 | 8.3 | 27.9 | 15.9 |
| **mmu-miR-744** | 5' | 1 | -2.44 | 0.17 | 0.028705 | -0.22 | 120.9 | 203.2 | 6.2 | 7.5 | 221.5 | 157.7 |
| **mmu-miR-30e*** | 3' | 1 | -1.18 | -0.92 | 0.032833 | 0.38 | 17.0 | 20.6 | 7.1 | 9.8 | 13.0 | 4.7 |
| **mmu-miR-302a** | 3' | 1 | 2.95 | -0.50 | 0.058518 | -1.41 | 9.5 | 12.7 | 48.6 | 105.7 | 6.3 | 9.9 |
| **mmu-miR-302a*** | 5' | 3 | -1.99 | -0.77 | 0.062322 | -1.41 | 1.4 | 1.7 | 0.3 | 0.5 | 0.5 | 1.9 |
| **mmu-miR-22** | 3' | 1 | -4.28 | -0.71 | 0.067302 | 3.12 | 194.5 | 191.4 | 11.1 | 9.8 | 131.7 | 61.9 |
| **mmu-let-7e** | 5' | -1 | -4.56 | -3.84 | 0.082897 | -2.82 | 649.5 | 991.9 | 4.5 | 17.6 | 97.9 | 5.9 |
| **mmu-miR-339-3p** | 3' | 2 | 2.91 | -1.36 | 0.084745 | -1.04 | 4.4 | 2.4 | 13.2 | 10.6 | 0.9 | 1.9 |
| **mmu-miR-136** | 5' | -3 | -3.17 | -0.97 | 0.099222 | -1.15 | 48.5 | 34.9 | 5.1 | 3.7 | 203.6 | 2.1 |
| **mmu-miR-181c*** | 3' | -1 | -1.93 | -1.37 | 0.103026 | 0.22 | 6.0 | 6.8 | 1.5 | 2.1 | 1.8 | 3.9 |
| **mmu-miR-154*** | 3' | 1 | -3.14 | 0.19 | 0.130709 | 1.14 | 64.3 | 54.4 | 12.9 | 4.4 | 136.8 | 12.2 |
| **mmu-miR-370** | 3' | -1 | -1.88 | -0.64 | 0.140638 | 1.19 | 18.7 | 40.5 | 3.6 | 9.8 | 21.2 | 18.7 |
| **mmu-miR-136*** | 3' | -1 | 1.23 | -2.60 | 0.143463 | -1.15 | 1.8 | 3.0 | 3.7 | 2.7 | 0.9 | 0.2 |
| **mmu-miR-151-3p** | 3' | 1 | -0.95 | 0.30 | 0.179627 | 0.84 | 44.8 | 80.0 | 30.7 | 34.4 | 69.5 | 46.0 |
| **mmu-miR-21*** | 3' | 1 | -1.05 | 0.09 | 0.181078 | -0.63 | 12.5 | 8.2 | 4.5 | 4.0 | 8.2 | 24.8 |
| **mmu-miR-127*** | 5' | 2 | 0.48 | -1.08 | 0.181543 | 0.50 | 2.8 | 3.2 | 5.6 | 3.0 | 2.4 | 0.9 |
| **mmu-miR-339-5p** | 5' | 1 | -1.36 | 0.70 | 0.190262 | -1.04 | 87.8 | 80.2 | 37.4 | 37.3 | 110.5 | 14.7 |
| **mmu-miR-370*** | 5' | -4 | 3.09 | 0.92 | 0.193099 | 1.19 | 1.9 | 2.9 | 4.0 | 10.6 | 2.3 | 6.8 |
| **mmu-miR-382*** | 3' | 1 | -0.86 | 0.06 | 0.220005 | -0.09 | 27.1 | 19.5 | 15.6 | 5.9 | 39.2 | 16.6 |
| **mmu-miR-301a** | 3' | 1 | -0.74 | -0.41 | 0.222864 | -0.53 | 89.2 | 120.4 | 45.3 | 52.0 | 66.7 | 39.6 |
| **mmu-miR-294** | 3' | 1 | 2.81 | -0.46 | 0.227992 | -0.27 | 101.1 | 62.6 | 626.9 | 486.8 | 75.3 | 43.1 |
| **mmu-miR-183*** | 3' | -1 | -0.69 | -0.47 | 0.258795 | 0.30 | 11.1 | 11.7 | 6.7 | 6.5 | 6.3 | 7.5 |
| **mmu-miR-674*** | 3' | 1 | -1.38 | 0.61 | 0.294604 | 0.21 | 24.8 | 48.0 | 15.1 | 16.9 | 33.2 | 27.4 |
| **mmu-miR-293*** | 5' | 1 | 2.34 | 0.36 | 0.30142 | 2.23 | 200.6 | 77.5 | 174.6 | 101.8 | 426.7 | 60.4 |
| **mmu-miR-376a*** | 5' | -1 | 1.58 | 0.41 | 0.31419 | -1.82 | 3.4 | 4.5 | 3.2 | 4.2 | 8.1 | 2.6 |
| **mmu-miR-26b** | 5' | 1 | -3.63 | -2.17 | 0.337873 | -2.43 | 2090.1 | 2392.2 | 42.5 | 116.2 | 255.6 | 273.4 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **mmu-miR-411*** | 3' | -2 | 3.72 | -0.69 | 0.357075 | -0.31 | 204.5 | 682.5 | 387.3 | 552.5 | 231.6 | 128.9 |
| **mmu-miR-152** | 3' | -1 | 5.14 | 0.26 | 0.39966 | 0.51 | 289.9 | 229.7 | 715.6 | 1621.5 | 34.0 | 171.2 |
| **mmu-miR-106b*** | 3' | -2 | 1.14 | -0.30 | 0.405039 | 0.23 | 4.3 | 9.6 | 7.2 | 15.0 | 4.9 | 5.9 |
| **mmu-miR-34c** | 5' | 3 | -4.34 | -1.87 | 0.440641 | -0.05 | 976.5 | 366.5 | 27.8 | 14.2 | 17.3 | 2503.3 |
| **mmu-miR-140*** | 3' | 1 | 1.96 | -1.75 | 0.47246 | -2.31 | 2.6 | 2.6 | 7.1 | 14.1 | 0.3 | 2.0 |
| **mmu-miR-186*** | 3' | 1 | 2.39 | -0.35 | 0.520653 | -0.32 | 2.2 | 2.9 | 8.7 | 4.0 | 1.6 | 1.9 |
| **mmu-miR-210** | 3' | -1 | -1.08 | -0.31 | 0.611119 | 0.85 | 31.1 | 20.4 | 19.0 | 7.4 | 27.9 | 13.7 |
| **mmu-miR-200b** | 3' | 1 | 1.84 | 1.71 | 0.611441 | -0.37 | 23.5 | 20.7 | 92.1 | 66.6 | 96.4 | 47.6 |
| **mmu-miR-140** | 5' | 1 | -1.02 | -0.18 | 0.614936 | -2.31 | 88.2 | 97.0 | 60.0 | 34.2 | 53.6 | 62.2 |
| **mmu-miR-669c** | 5' | 1 | -1.80 | 1.49 | 0.616742 | -0.96 | 33.0 | 60.5 | 11.2 | 17.8 | 125.0 | 204.2 |
| **mmu-miR-301b** | 3' | 1 | 1.29 | 0.19 | 0.622734 | 0.45 | 17.6 | 24.4 | 31.7 | 55.6 | 16.7 | 25.5 |
| **mmu-miR-148a*** | 5' | -1 | 0.52 | 0.64 | 6.42E-01 | 1.76 | 3.8 | 5.4 | 4.9 | 8.4 | 4.9 | 8.7 |
| **mmu-miR-20a** | 5' | 1 | -1.45 | -0.23 | 0.686413 | 0.88 | 360.4 | 588.3 | 192.5 | 130.3 | 562.2 | 257.9 |
| **mmu-miR-674** | 5' | -1 | -1.35 | -1.27 | 0.694002 | 0.21 | 7.7 | 12.7 | 4.5 | 3.8 | 6.3 | 2.8 |
| **mmu-miR-294*** | 5' | -1 | 1.92 | -0.92 | 0.711832 | -0.27 | 30.6 | 15.0 | 90.1 | 60.4 | 32.1 | 3.8 |
| **mmu-miR-540-3p** | 3' | -1 | -0.67 | -0.31 | 0.725407 | 0.55 | 90.6 | 191.2 | 68.4 | 124.1 | 85.9 | 93.6 |
| **mmu-miR-96*** | 3' | 1 | -1.60 | 0.11 | 0.728546 | 0.59 | 4.9 | 10.9 | 3.3 | 5.6 | 9.5 | 4.5 |
| **mmu-miR-93** | 5' | 1 | -1.37 | -0.82 | 0.731318 | -1.37 | 188.7 | 271.4 | 63.2 | 77.4 | 102.2 | 166.2 |
| **mmu-miR-295*** | 5' | -1 | 1.10 | -0.03 | 0.743196 | 0.65 | 24.0 | 13.3 | 52.1 | 29.0 | 12.0 | 17.2 |
| **mmu-miR-130a** | 3' | 1 | 1.42 | 1.99 | 0.750725 | -1.26 | 52.4 | 71.6 | 150.7 | 178.0 | 143.5 | 342.2 |
| **mmu-miR-361** | 5' | -1 | -1.00 | -3.55 | 0.75438 | -5.21 | 506.1 | 614.3 | 96.9 | 206.8 | 27.6 | 81.0 |
| **mmu-miR-291a-3p** | 3' | 1 | 0.98 | 1.44 | 0.775377 | 0.79 | 3.4 | 1.9 | 5.9 | 4.2 | 13.9 | 3.4 |
| **mmu-miR-290-3p** | 3' | 1 | 3.07 | -0.43 | 0.77542 | -1.60 | 1.6 | 1.9 | 8.5 | 18.5 | 1.4 | 1.2 |
| **mmu-miR-181c** | 5' | -1 | 1.31 | 1.19 | 0.783583 | 0.22 | 57.1 | 41.3 | 25.4 | 31.8 | 193.3 | 56.4 |
| **mmu-miR-130b** | 3' | 1 | 1.31 | 0.03 | 0.785233 | 0.75 | 68.0 | 89.1 | 124.4 | 281.8 | 37.5 | 166.7 |
| **mmu-miR-363-5p** | 5' | 1 | -1.28 | -0.34 | 0.792182 | -0.43 | 30.6 | 14.3 | 18.3 | 8.2 | 28.8 | 80.1 |
| **mmu-miR-146b** | 5' | 1 | -2.35 | -2.59 | 0.814368 | -3.40 | 315.3 | 332.1 | 12.9 | 34.3 | 20.8 | 225.4 |
| **mmu-miR-130b*** | 5' | -1 | 1.42 | -0.02 | 0.81995 | 0.75 | 26.1 | 36.2 | 81.8 | 39.3 | 60.3 | 9.8 |
| **mmu-miR-342-5p** | 5' | 1 | -0.66 | -0.91 | 0.820067 | -0.26 | 8.8 | 4.5 | 3.7 | 4.0 | 3.5 | 4.2 |
| **mmu-miR-322*** | 3' | 2 | -0.18 | -0.62 | 0.850264 | 0.11 | 1.5 | 3.4 | 1.2 | 3.7 | 1.8 | 0.9 |
| **mmu-miR-191*** | 3' | 1 | -0.45 | -1.55 | 0.857033 | -0.52 | 1.8 | 5.0 | 1.2 | 3.5 | 2.3 | 0.7 |
| **mmu-miR-125a-3p** | 3' | 1 | -0.50 | -0.21 | 0.867815 | 1.36 | 5.4 | 5.8 | 3.2 | 4.3 | 4.6 | 3.6 |
| **mmu-miR-541** | 5' | 1 | -1.71 | -0.20 | 0.876335 | -1.38 | 201.0 | 184.0 | 56.6 | 63.2 | 165.3 | 184.9 |
| **mmu-miR-181d** | 5' | -1 | 1.64 | 0.35 | 0.879061 | 1.01 | 188.6 | 287.8 | 42.1 | 132.0 | 260.6 | 342.0 |
| **mmu-miR-148b** | 3' | 1 | -0.31 | 0.03 | 0.879146 | -0.43 | 502.0 | 464.0 | 594.3 | 750.9 | 360.9 | 1535.5 |
| **mmu-miR-302d*** | 5' | -1 | 0.17 | -0.50 | 0.880356 | -1.35 | 1.6 | 1.8 | 1.8 | 1.9 | 2.1 | 0.9 |
| **mmu-miR-872** | 5' | 1 | -0.05 | -1.53 | 0.882873 | -0.88 | 638.5 | 743.0 | 87.3 | 190.5 | 81.9 | 510.9 |
| **mmu-miR-431** | 5' | 2 | -1.65 | 0.73 | 0.893341 | 0.44 | 103.3 | 192.1 | 6.0 | 21.5 | 212.1 | 40.6 |
| **mmu-miR-382** | 5' | 1 | -0.21 | 0.03 | 0.894186 | -0.09 | 8.9 | 17.5 | 7.7 | 12.9 | 8.2 | 19.7 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **mmu-miR-495** | 3' | 1 | -0.75 | 0.40 | 0.894688 | -0.32 | 10.9 | 11.0 | 6.5 | 6.5 | 15.9 | 7.2 |
| **mmu-miR-33** | 5' | -4 | -0.74 | -1.34 | 0.898585 | -0.79 | 256.1 | 209.1 | 163.6 | 72.0 | 716.9 | 22.1 |
| **mmu-miR-299*** | 5' | -1 | 1.20 | 1.89 | 0.899385 | 7.60 | 3.2 | 2.3 | 6.3 | 6.0 | 25.0 | 3.6 |
| **mmu-miR-17** | 5' | -1 | 0.01 | -0.66 | 0.902599 | -0.06 | 87.7 | 79.8 | 67.9 | 104.0 | 28.2 | 96.4 |
| **mmu-miR-429** | 3' | -1 | -2.12 | 0.13 | 0.90313 | 0.41 | 194.5 | 189.8 | 50.2 | 21.6 | 232.2 | 172.7 |
| **mmu-miR-134** | 5' | -1 | -1.49 | 0.58 | 0.919668 | 0.23 | 334.3 | 246.7 | 158.4 | 50.6 | 140.5 | 670.1 |
| **mmu-miR-125a-5p** | 5' | 1 | 0.45 | 1.48 | 0.922301 | 1.36 | 40.6 | 334.7 | 81.2 | 177.3 | 252.4 | 43.6 |
| **mmu-miR-423-5p** | 5' | 2 | 1.10 | -2.84 | 0.922494 | 2.67 | 61.8 | 38.6 | 47.7 | 39.2 | 26.8 | 65.2 |
| **mmu-miR-425** | 5' | 1 | -0.47 | 1.10 | 0.929207 | -0.69 | 35.5 | 31.6 | 17.1 | 12.7 | 51.3 | 37.9 |
| **mmu-miR-106a** | 5' | 1 | 1.18 | -0.66 | 0.929272 | -2.69 | 31.3 | 34.3 | 55.1 | 56.5 | 23.7 | 18.3 |
| **mmu-miR-200b*** | 5' | 1 | -0.46 | -0.42 | 0.929375 | -0.37 | 10.0 | 11.9 | 7.7 | 9.0 | 6.2 | 12.7 |
| **mmu-miR-148a** | 3' | 1 | 0.75 | 0.08 | 0.931069 | 1.76 | 1237.6 | 1137.8 | 959.4 | 893.0 | 609.7 | 2140.5 |
| **mmu-miR-322** | 5' | -1 | 0.11 | 0.08 | 0.933339 | 0.11 | 27.9 | 130.0 | 36.2 | 117.4 | 26.0 | 34.1 |
| **mmu-miR-503** | 5' | 1 | -1.72 | -1.76 | 0.94047 | -0.42 | 400.1 | 249.0 | 139.0 | 102.7 | 17.9 | 145.6 |
| **mmu-miR-434-3p** | 3' | -1 | -0.91 | 1.39 | 0.940481 | 1.01 | 20.7 | 16.3 | 9.9 | 10.1 | 29.5 | 75.4 |
| **mmu-miR-363-3p** | 3' | 1 | 0.60 | -1.37 | 0.941846 | -0.43 | 2.8 | 2.8 | 3.5 | 4.9 | 1.2 | 1.0 |
| **mmu-miR-21** | 5' | -1 | 0.02 | -0.18 | 0.947997 | -0.63 | 317.8 | 346.9 | 452.6 | 212.3 | 220.6 | 391.7 |
| **mmu-miR-195** | 5' | 3 | 0.16 | 0.32 | 0.958569 | -0.79 | 270.8 | 999.1 | 180.3 | 1234.9 | 207.6 | 570.0 |
| **mmu-miR-543** | 3' | -2 | 3.05 | 1.78 | 0.958896 | -0.72 | 156.2 | 72.5 | 193.6 | 69.8 | 787.5 | 15.9 |
| **mmu-miR-22*** | 5' | 1 | 0.07 | -1.25 | 0.96206 | 3.12 | 37.1 | 22.3 | 41.0 | 14.6 | 36.6 | 7.7 |
| **mmu-miR-34a** | 5' | -2 | 0.23 | -0.43 | 0.967341 | 2.01 | 237.2 | 320.0 | 155.8 | 58.6 | 446.6 | 119.5 |
| **mmu-miR-222** | 3' | -1 | -0.89 | -0.74 | 0.969783 | -0.66 | 50.8 | 59.2 | 27.5 | 37.2 | 46.1 | 13.8 |
| **mmu-miR-411** | 5' | -1 | -0.21 | 0.24 | 0.971829 | -0.31 | 4.1 | 4.0 | 4.2 | 2.9 | 3.2 | 7.0 |
| **mmu-miR-20b** | 5' | 1 | -0.72 | -1.10 | 0.977787 | -0.80 | 148.3 | 399.8 | 97.8 | 204.4 | 71.5 | 198.3 |
| **mmu-miR-15a** | 5' | 1 | 0.26 | 0.47 | 0.978257 | -1.63 | 231.7 | 206.8 | 46.9 | 24.4 | 216.2 | 182.5 |
| **mmu-miR-151-5p** | 5' | -1 | -0.62 | -0.06 | 0.978586 | 0.84 | 128.7 | 143.1 | 85.7 | 110.0 | 116.8 | 85.9 |
| **mmu-miR-376a** | 3' | -1 | 1.19 | -0.09 | 0.981146 | -1.82 | 96.3 | 113.7 | 82.4 | 141.9 | 129.9 | 44.7 |
| **mmu-miR-32** | 5' | -1 | -0.41 | 1.03 | 0.981738 | 0.68 | 59.2 | 46.4 | 49.6 | 29.5 | 109.5 | 51.2 |
| **mmu-miR-154** | 5' | 1 | -0.61 | -1.32 | 0.981928 | 1.14 | 51.8 | 36.6 | 61.9 | 27.8 | 106.8 | 9.7 |
| **mmu-miR-499** | 5' | 1 | -0.65 | 1.96 | 0.98236 | -0.11 | 23.1 | 12.9 | 7.7 | 10.9 | 24.1 | 4.7 |
| **mmu-miR-148b*** | 5' | -1 | 0.01 | -1.63 | 0.982613 | -0.43 | 23.8 | 24.5 | 29.3 | 27.9 | 3.7 | 37.5 |
| **mmu-miR-296-5p** | 5' | NA | -0.63 | -0.14 | 0.984009 | 6.30 | 30.7 | 14.1 | 35.0 | 6.5 | 214.1 | 6.3 |
| **mmu-miR-186** | 5' | 1 | 0.70 | -1.42 | 0.984462 | -0.32 | 22.6 | 38.3 | 34.6 | 67.0 | 12.0 | 10.1 |
| **mmu-miR-183** | 5' | 1 | 0.07 | -0.03 | 0.984776 | 0.30 | 4.3 | 4.1 | 3.6 | 5.4 | 3.4 | 5.0 |
| **mmu-miR-30a** | 5' | 1 | -0.69 | -1.45 | 0.990709 | 0.05 | 153.9 | 155.3 | 89.3 | 82.9 | 25.7 | 110.0 |
| **mmu-miR-150** | 5' | 1 | -2.19 | 0.32 | 0.991442 | -0.71 | 269.4 | 175.6 | 56.9 | 34.2 | 222.8 | 30.6 |
| **mmu-miR-673-5p** | 5' | 1 | -0.52 | -1.25 | 0.991521 | -1.35 | 14.3 | 18.4 | 7.0 | 16.2 | 11.1 | 4.2 |
| **mmu-miR-20b*** | 3' | -1 | 0.85 | 2.56 | 0.993301 | -0.80 | 37.4 | 21.2 | 46.4 | 14.2 | 133.2 | 9.9 |
| **mmu-miR-15b** | 5' | 3 | -1.54 | -2.97 | 0.993358 | -0.57 | 916.8 | 5269.2 | 654.5 | 3491.3 | 228.5 | 457.4 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **mmu-miR-182*** | 3' | -5 | -0.93 | 5.17 | 0.993368 | 0.30 | 0.6 | 2.9 | 0.6 | 0.9 | 9.9 | 5.2 |
| **mmu-miR-28** | 5' | -1 | 0.92 | -0.54 | 0.995203 | -0.32 | 217.4 | 198.4 | 141.8 | 460.3 | 56.9 | 213.9 |
| **mmu-miR-672** | 5' | 1 | 0.26 | -0.74 | 0.995529 | 0.49 | 422.5 | 419.7 | 543.9 | 533.6 | 138.0 | 383.0 |
| **mmu-miR-191** | 5' | 1 | -0.07 | -0.43 | 0.995672 | -0.52 | 41.4 | 53.6 | 42.7 | 45.0 | 46.6 | 26.0 |
| **mmu-miR-101a*** | 5' | 1 | -0.55 | -1.85 | 0.996444 | -0.49 | 1.4 | 1.3 | 0.9 | 1.0 | 0.7 | 0.5 |
| **mmu-miR-292-5p** | 5' | 1 | -1.84 | -1.82 | 0.997362 | -0.33 | 1487.5 | 2754.2 | 566.8 | 660.6 | 1145.5 | 379.2 |
| **mmu-miR-369-5p** | 5' | 3 | -1.65 | -4.53 | 0.997465 | -1.07 | 398.3 | 461.7 | 27.9 | 24.9 | 24.5 | 29.2 |
| **mmu-let-7g** | 5' | -1 | 0.59 | 0.64 | 0.99756 | -2.34 | 203.7 | 121.3 | 75.8 | 214.5 | 52.9 | 184.1 |
| **mmu-miR-96** | 5' | 1 | -0.33 | -1.70 | 0.99787 | 0.59 | 852.7 | 1075.7 | 660.5 | 1012.9 | 310.5 | 151.2 |
| **mmu-miR-379*** | 3' | -1 | 2.12 | -1.76 | 0.998424 | -0.38 | 233.0 | 263.8 | 187.2 | 254.8 | 123.4 | 37.5 |
| **mmu-miR-30e** | 5' | 1 | -0.50 | -1.09 | 0.998807 | 0.38 | 124.6 | 110.9 | 73.0 | 97.7 | 35.2 | 85.3 |
| **mmu-miR-291a-5p** | 5' | 1 | -0.58 | 0.75 | 0.998851 | 0.79 | 66.0 | 70.7 | 47.9 | 43.9 | 84.2 | 145.5 |
| **mmu-miR-182** | 5' | 1 | -0.39 | -1.25 | 0.998959 | 0.30 | 826.7 | 938.4 | 552.1 | 850.0 | 303.0 | 296.3 |
| **mmu-miR-126-5p** | 5' | 1 | 0.05 | -0.27 | 0.999125 | 5.10 | 94.6 | 114.0 | 131.9 | 115.6 | 23.3 | 35.7 |
| **mmu-miR-708** | 5' | 1 | -0.08 | 0.10 | 0.99921 | -0.31 | 68.0 | 64.8 | 59.0 | 65.4 | 39.7 | 110.2 |
| **mmu-miR-30b** | 5' | -1 | 1.38 | -1.16 | 0.999368 | -0.10 | 708.0 | 551.9 | 499.9 | 646.5 | 192.3 | 76.3 |
| **mmu-miR-135b** | 5' | -1 | -0.25 | 0.42 | 0.999408 | -0.06 | 499.4 | 354.8 | 474.4 | 357.6 | 411.3 | 113.4 |
| **mmu-miR-142-5p** | 5' | 2 | -0.07 | -0.28 | 0.99952 | 0.09 | 1.1 | 1.5 | 1.1 | 1.3 | 0.7 | 1.6 |
| **mmu-miR-143** | 3' | -1 | 1.60 | 0.09 | 0.999696 | -0.34 | 308.0 | 333.4 | 363.7 | 728.8 | 32.0 | 663.0 |
| **mmu-miR-484** | 5' | 1 | 0.56 | -1.06 | 0.999725 | 1.77 | 113.1 | 182.3 | 228.9 | 235.6 | 321.9 | 42.8 |
| **mmu-miR-379** | 5' | -1 | 0.04 | -0.57 | 0.99988 | -0.38 | 503.2 | 337.7 | 347.5 | 374.7 | 109.5 | 471.1 |
| **mmu-miR-106b** | 5' | 1 | -0.74 | -2.68 | 0.999917 | 0.23 | 1546.9 | 1169.6 | 840.6 | 788.3 | 187.0 | 271.5 |
| **mmu-miR-15b*** | 3' | 1 | 0.81 | -1.48 | 0.999917 | -0.57 | 80.8 | 71.9 | 82.1 | 80.3 | 25.0 | 12.1 |
| **mmu-miR-30d** | 5' | 1 | -0.61 | -3.52 | 0.999919 | -0.78 | 1018.1 | 1378.8 | 664.6 | 678.9 | 82.3 | 209.5 |
| **mmu-miR-3096-3p** | 3' | 1 | -0.14 | -2.29 | 0.99996 | 0.69 | 3.1 | 1.7 | 2.5 | 2.1 | 1.0 | 0.5 |
| **mmu-miR-146b*** | 3' | NA | -1.16 | -0.66 | NA | -3.40 | 1.1 | 1.2 | 1.1 | 1.0 | 1.5 | 1.0 |
| **mmu-let-7d*** | 3' | 1 | -2.97 | -3.18 | NA | -3.24 | 2.1 | 1.9 | 1.1 | 1.0 | 1.2 | 0.9 |
| **mmu-miR-98** | 5' | 3 | -4.62 | -4.79 | NA | -3.20 | 154.2 | 201.4 | 1.6 | 5.8 | 7.1 | 16.9 |
| **mmu-miR-98*** | 3' | 1 | -1.09 | 0.49 | NA | -3.20 | 1.5 | 0.9 | 0.9 | 1.0 | 1.0 | 1.1 |
| **mmu-let-7e*** | 3' | 1 | -4.10 | -0.94 | NA | -2.82 | 1.7 | 2.4 | 0.9 | 1.0 | 3.6 | 1.2 |
| **mmu-miR-106a*** | 3' | -1 | -1.79 | 3.64 | NA | -2.69 | 1.0 | 1.2 | 0.9 | 1.0 | 12.6 | 2.4 |
| **mmu-miR-26b*** | 3' | -2 | -1.76 | 0.42 | NA | -2.43 | 3.3 | 2.5 | 1.2 | 1.0 | 6.6 | 1.2 |
| **mmu-let-7g*** | 3' | -1 | -0.43 | -1.22 | NA | -2.34 | 1.2 | 1.5 | 1.1 | 1.0 | 1.0 | 1.0 |
| **mmu-miR-185*** | 3' | -1 | -1.40 | 1.58 | NA | -2.22 | 1.2 | 1.8 | 1.1 | 1.0 | 2.4 | 2.3 |
| **mmu-miR-300*** | 5' | 1 | -1.64 | -1.45 | NA | -2.06 | 2.0 | 2.0 | 0.8 | 1.0 | 1.0 | 0.9 |
| **mmu-miR-15a*** | 3' | 1 | -3.26 | -1.85 | NA | -1.63 | 3.6 | 4.4 | 1.3 | 1.2 | 8.5 | 1.4 |
| **mmu-miR-93*** | 3' | -1 | -2.43 | 2.03 | NA | -1.37 | 10.0 | 11.8 | 2.5 | 1.7 | 30.8 | 7.0 |
| **mmu-miR-673-3p** | 3' | -1 | 1.39 | 1.70 | NA | -1.35 | 2.2 | 4.8 | 3.2 | 3.7 | 4.8 | 8.2 |
| **mmu-miR-130a*** | 5' | -1 | 0.60 | -0.35 | NA | -1.26 | 7.8 | 7.4 | 7.8 | 5.9 | 19.5 | 2.0 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **mmu-miR-99b*** | 3' | 1 | -4.19 | 0.29 | NA | -1.06 | 2.9 | 6.2 | 1.1 | 1.0 | 5.5 | 17.5 |
| **mmu-miR-99b** | 5' | 1 | -1.06 | -0.21 | NA | -1.06 | 52.5 | 66.9 | 3.2 | 5.0 | 26.2 | 98.7 |
| **mmu-miR-669c*** | 3' | NA | -0.79 | -0.79 | NA | -0.96 | 1.0 | 1.1 | 1.0 | 1.0 | 1.0 | 1.0 |
| **mmu-miR-302b*** | 5' | NA | 0.20 | 0.93 | NA | -0.82 | 2.8 | 1.7 | 4.2 | 1.6 | 10.5 | 2.8 |
| **mmu-miR-33*** | 3' | 1 | -2.97 | 0.02 | NA | -0.79 | 3.7 | 4.3 | 0.5 | 1.8 | 5.9 | 1.1 |
| **mmu-miR-195*** | 3' | 1 | -1.09 | -2.09 | NA | -0.79 | 2.7 | 3.6 | 2.6 | 2.5 | 1.7 | 1.1 |
| **mmu-miR-30d*** | 3' | 1 | -4.21 | -0.32 | NA | -0.78 | 3.7 | 5.4 | 0.2 | 0.4 | 3.1 | 2.1 |
| **mmu-miR-543*** | 5' | -1 | -0.36 | -0.58 | NA | -0.72 | 1.0 | 1.1 | 0.8 | 1.0 | 0.8 | 0.9 |
| **mmu-miR-150*** | 3' | -1 | 0.06 | 0.53 | NA | -0.71 | 0.8 | 2.7 | 1.3 | 1.1 | 0.7 | 4.4 |
| **mmu-miR-425*** | 3' | 1 | -0.59 | -0.44 | NA | -0.69 | 1.2 | 1.0 | 1.2 | 0.9 | 0.6 | 1.2 |
| **mmu-miR-222*** | 5' | 2 | -0.73 | -2.32 | NA | -0.66 | 3.6 | 1.2 | 1.4 | 1.5 | 0.7 | 1.0 |
| **mmu-miR-301a*** | 5' | 1 | -3.18 | 0.44 | NA | -0.53 | 3.1 | 3.6 | 1.2 | 1.3 | 17.6 | 2.2 |
| **mmu-miR-19a*** | 5' | 1 | 0.08 | -0.74 | NA | -0.48 | 1.8 | 2.0 | 2.7 | 2.2 | 2.9 | 1.0 |
| **mmu-miR-122** | 5' | 2 | 0.55 | -0.38 | NA | -0.38 | 262.9 | 2.6 | 361.9 | 7.4 | 7.5 | 19.0 |
| **mmu-miR-122*** | 3' | -2 | 0.53 | -0.64 | NA | -0.38 | 1.6 | 1.0 | 4.6 | 1.0 | 1.0 | 1.0 |
| **mmu-miR-143*** | 5' | -1 | -1.22 | -0.59 | NA | -0.34 | 2.5 | 2.9 | 1.8 | 1.2 | 3.5 | 1.2 |
| **mmu-miR-495*** | 5' | 1 | 0.67 | -1.29 | NA | -0.32 | 9.5 | 8.5 | 6.0 | 6.0 | 15.2 | 2.4 |
| **mmu-miR-744*** | 3' | -2 | -3.94 | -0.36 | NA | -0.22 | 5.3 | 6.9 | 1.1 | 1.1 | 6.9 | 2.0 |
| **mmu-miR-499*** | 3' | 1 | 0.25 | -2.35 | NA | -0.11 | 2.2 | 1.4 | 1.1 | 1.3 | 0.5 | 1.0 |
| **mmu-miR-34c*** | 3' | 1 | -6.02 | -0.62 | NA | -0.05 | 6.7 | 7.9 | 0.4 | 0.5 | 2.3 | 3.4 |
| **mmu-miR-380-3p** | 3' | -5 | -7.37 | -1.77 | NA | 0.00 | 358.0 | 546.8 | 2.8 | 3.3 | 310.8 | 100.8 |
| **mmu-miR-380-5p** | 5' | 1 | -4.81 | -2.26 | NA | 0.00 | 121.8 | 91.9 | 1.5 | 2.2 | 6.3 | 78.7 |
| **mmu-miR-741** | 3' | 2 | 4.60 | -0.09 | NA | 0.11 | 3.6 | 4.8 | 56.5 | 110.8 | 1.4 | 7.4 |
| **mmu-miR-741*** | 5' | NA | 0.67 | -4.71 | NA | 0.11 | 7.5 | 8.0 | 13.1 | 12.0 | 1.2 | 1.6 |
| **mmu-miR-320*** | 5' | 1 | -0.21 | 0.29 | NA | 0.17 | 0.9 | 1.0 | 0.9 | 1.0 | 1.0 | 1.0 |
| **mmu-miR-134*** | 3' | 1 | -2.61 | 0.61 | NA | 0.23 | 1.6 | 1.9 | 0.8 | 0.4 | 2.2 | 1.4 |
| **mmu-miR-429*** | 5' | NA | -2.58 | -2.15 | NA | 0.41 | 2.0 | 2.8 | 1.3 | 1.1 | 6.4 | 1.0 |
| **mmu-miR-431*** | 3' | 1 | 1.44 | -0.54 | NA | 0.44 | 1.2 | 1.5 | 1.1 | 1.4 | 0.8 | 1.2 |
| **mmu-miR-301b*** | 5' | NA | -0.50 | 1.23 | NA | 0.45 | 1.0 | 1.2 | 1.1 | 1.0 | 4.5 | 1.0 |
| **mmu-miR-672*** | 3' | 3 | 2.30 | -2.39 | NA | 0.49 | 17.6 | 12.4 | 31.0 | 15.0 | 4.2 | 1.9 |
| **mmu-miR-152*** | 5' | -1 | -1.49 | -1.20 | NA | 0.51 | 2.0 | 1.7 | 1.5 | 1.9 | 1.7 | 1.2 |
| **mmu-miR-540-5p** | 5' | 1 | 0.54 | 0.34 | NA | 0.55 | 3.7 | 3.6 | 5.5 | 4.0 | 14.3 | 4.1 |
| **mmu-miR-539-5p** | 5' | 1 | -4.50 | -2.67 | NA | 0.61 | 70.3 | 174.7 | 2.1 | 3.1 | 42.6 | 6.9 |
| **mmu-miR-539-3p** | 3' | 1 | -3.77 | -3.35 | NA | 0.61 | 26.1 | 37.1 | 1.3 | 1.6 | 1.2 | 2.6 |
| **mmu-miR-32*** | 3' | -1 | 3.98 | 0.56 | NA | 0.68 | 4.6 | 7.4 | 9.7 | 6.1 | 2.2 | 4.1 |
| **mmu-miR-3096-5p** | 5' | -1 | 0.01 | -3.29 | NA | 0.69 | 3.6 | 4.4 | 3.6 | 4.6 | 0.3 | 1.1 |
| **mmu-miR-210*** | 5' | 1 | -0.27 | -0.82 | NA | 0.85 | 3.2 | 3.5 | 3.1 | 2.5 | 6.4 | 1.4 |
| **mmu-miR-216b** | 5' | 1 | -0.38 | -2.85 | NA | 0.95 | 1.1 | 201.9 | 1.1 | 216.7 | 1.5 | 1.2 |
| **mmu-miR-216b*** | 3' | NA | 0.00 | 0.00 | NA | 0.95 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **mmu-miR-181d*** | 3' | 1 | -4.02 | -0.53 | NA | 1.01 | 1.5 | 0.9 | 0.7 | 0.6 | 1.1 | 1.1 |
| **mmu-miR-337-3p** | 3' | -1 | -4.28 | -0.71 | NA | 1.06 | 2.7 | 1.6 | 0.4 | 0.5 | 3.6 | 0.6 |
| **mmu-miR-101b*** | 5' | -1 | -3.45 | 0.26 | NA | 1.20 | 1.5 | 1.0 | 0.2 | 0.4 | 1.2 | 1.1 |
| **(mmu-miR-484-3p)** | 3' | -5 | -0.71 | 0.95 | NA | 1.77 | 1.0 | 1.0 | 1.0 | 0.9 | 1.0 | 1.3 |
| **mmu-miR-25*** | 5' | 2 | -5.47 | 0.16 | NA | 1.93 | 26.9 | 12.7 | 1.2 | 1.1 | 3.0 | 1212.6 |
| **mmu-miR-34a*** | 3' | -1 | -0.94 | -0.09 | NA | 2.01 | 2.4 | 1.0 | 1.0 | 1.0 | 2.3 | 1.1 |
| **mmu-miR-410*** | 5' | 6 | -2.76 | 0.14 | NA | 2.19 | 11.1 | 30.4 | 2.8 | 3.8 | 14.4 | 36.2 |
| **mmu-miR-340-3p** | 3' | -1 | -4.74 | -0.92 | NA | 2.58 | 7.7 | 10.5 | 0.9 | 1.0 | 3.2 | 2.6 |
| **mmu-miR-200c*** | 5' | 1 | -0.75 | -1.36 | NA | 2.93 | 8.4 | 8.5 | 7.0 | 6.0 | 3.5 | 2.2 |
| **mmu-miR-381*** | 5' | 2 | -1.80 | 0.30 | NA | 3.03 | 2.0 | 5.2 | 1.5 | 2.0 | 26.1 | 1.1 |
| **mmu-miR-378*** | 5' | 1 | 2.71 | -0.98 | NA | 3.46 | 3.2 | 3.4 | 8.0 | 6.9 | 3.9 | 1.5 |
| **mmu-miR-377** | 3' | -1 | -1.05 | -0.01 | NA | 3.74 | 7.8 | 8.5 | 1.8 | 1.3 | 8.9 | 3.7 |
| **mmu-miR-377*** | 5' | 2 | 0.75 | 2.25 | NA | 3.74 | 2.4 | 1.9 | 1.1 | 1.9 | 11.9 | 3.0 |
| **mmu-miR-494*** | 5' | 3 | -0.61 | 0.16 | NA | 3.84 | 0.9 | 2.8 | 0.9 | 1.3 | 0.5 | 2.8 |
| **mmu-miR-29a*** | 5' | 2 | 1.87 | -0.94 | NA | 4.15 | 6.0 | 5.4 | 28.3 | 7.0 | 6.0 | 1.0 |
| **mmu-miR-496*** | 5' | 1 | 0.21 | 0.26 | NA | 4.42 | 1.0 | 1.1 | 1.0 | 1.1 | 1.7 | 1.0 |
| **(mmu-miR-1983-5p)** | 5' | -11 | 2.31 | -0.91 | NA | 7.77 | 10.0 | 12.4 | 24.5 | 29.6 | 3.3 | 3.1 |

## II-5. Crystal structure of human Dicer PAZ-siRNA complex

To study the detailed mechanism of 5′-phosphate recognition by Dicer, I collaborated with Dr. Dinshaw Patels' group (Yuan Tian, Dhirendra K. Simanshu and Jin-Biao Ma) to solve the crystal structure of human Dicer PAZ cassette bound to siRNA. Human Dicer construct used for structure analysis (amino acid 755~1055) included PAZ domain with flanking platform domain and connector helix domain which were previously discovered by structural study on *G. intestinalis* Dicer (**Figure II-18A**). Tian et al. first used 12-mer siRNA duplex for co-crystallization with Dicer construct and the structure of Dicer-siRNA complex was refined to 1.95Å resolution (**Figure II-18D** and **Figure II-18E**). Overall, structure of human Dicer PAZ cassette largely overlapped to what has been observed in *G. intestinalis* Dicer. Especially, 2nt 3′-overhang was inserted into 3′-pocket within Dicer PAZ domain, and anchored by a multitude of intermolecular hydrogen bonds (Tyr926, Tyr961, Tyr 962, Tyr966 and Arg927), consistent with previous studies on PAZ domains of *G. intestinalis* Dicer or Argonautes (**Figure II-19A**). When human Dicer mutated at these residues are tested, 3′ recognition pattern was disturbed, validating this structural finding (**Figure II-8B**).

**Figure II-18. Dicer domain architecture, species-specific sequences of 'Dicer PAZ insertion element' and structure of hDicer PAZ cassette bound to a 12-mer siRNA[8]**

(A) Domain architecture of hDicer (top) and hDicer 'platform-PAZ-connector helix' cassette (below), with the latter used for structural studies of complexes with bound siRNAs and dsRNA. The Dicer-specific insert is shown in red.

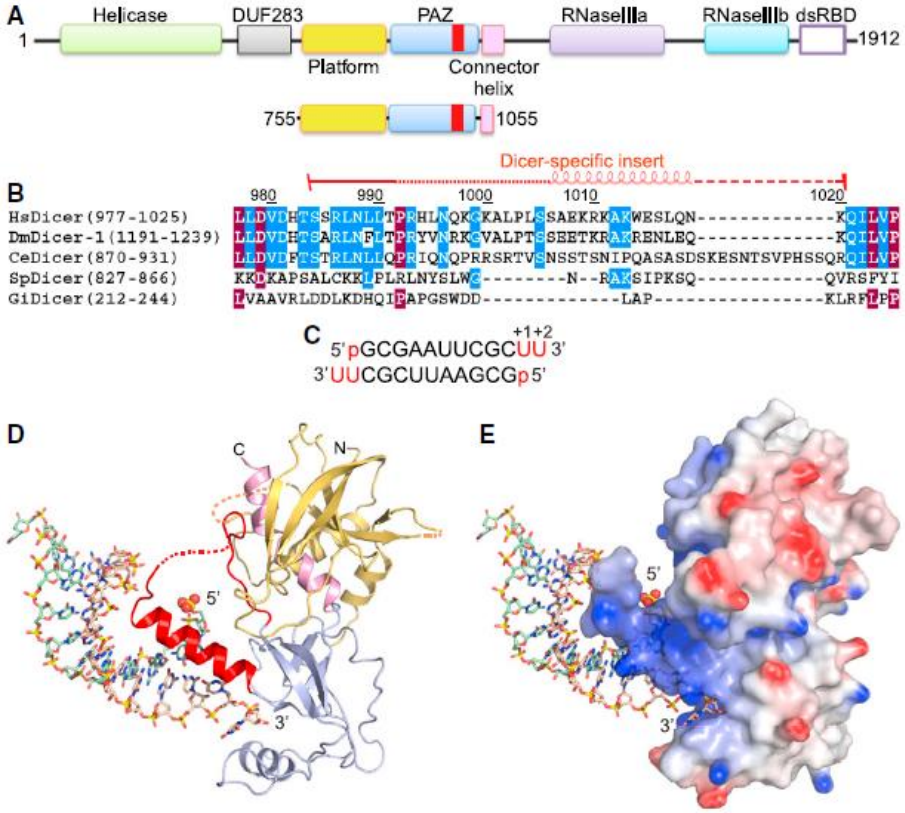
(B) Comparison of sequences (human, D. melanogaster-1, C. elegans, S. pombe and G. intestinalis) of the 'Dicer PAZ insert' element.

(C) Sequence of the self-complementary 12-mer siRNA containing 5′-p and UU-overhangs at 3′-ends.

(D) Structure of hDicer PAZ cassette bound to a 12-mer siRNA. The platform, PAZ and connector helix are colored in yellow, blue and pink, respectively. The 'Dicer PAZ-insertion element' is colored in red and composed of a disordered segment followed by the 'hDicer-specific helix'. The 5′-phosphate is shown in a space-filling representation. (E) A view of the complex in the same orientation as in panel D, with the protein in an electrostatic surface representation.

---

[8] Tian, Y., Simanshu, D.K., Ma, J.B., Park, J.E., Heo, I., Kim, V.N., and Patel, D.J. (2014). A phosphate-binding pocket within the platform-PAZ-connector helix cassette of human Dicer. Molecular cell *53*, 606-616.


A
Helicase
DUF283
PAZ
RNaseIIIa
RNaseIIIb
dsRBD
1
1912
Platform
Connector helix
755
1055
Dicer-specific insert
B
980
990
1000
1010
1020
HsDicer(977-1025)
DmDicer-1(1191-1239)
CeDicer(870-931)
SpDicer(827-866)
GiDicer(212-244)
C
5'pGCGAAUUCGCUU3'
+1+2
3'UUCGCUUAAGCGp5'
D
C
N
5'
3'
E
5'
3'

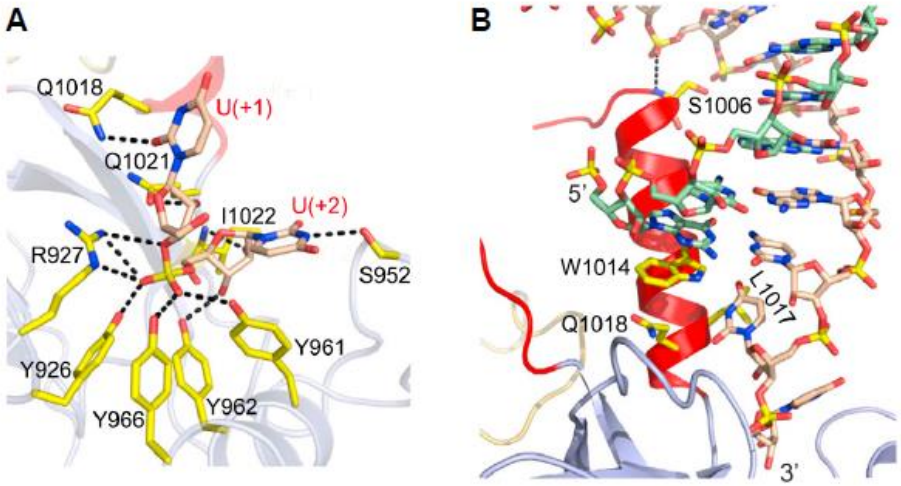



**Figure II-19. Structural details of hDicer PAZ cassette bound to 12mer siRNA[9]**

(A) Intermolecular hydrogen bonds involving the backbone phosphate and 2′-OH groups of the 2 nt overhang at the 3′ end of the 12-mer siRNA and residues lining the 3′-pocket in the complex

(B) Intermolecular hydrogen-bonding contacts involving either end of the hDicer-specific helix, with Trp1014 stacking over the terminal base pair.

[9] Ibid.

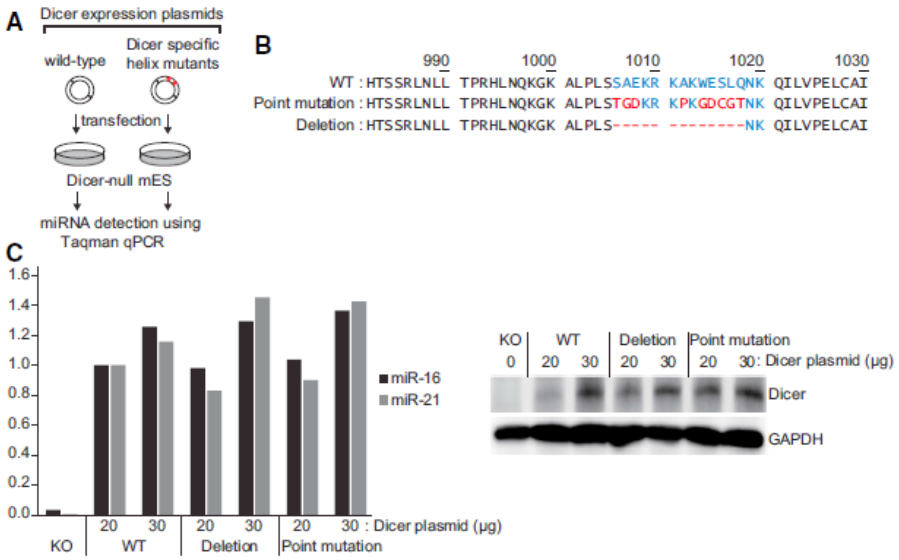
**Figure II-20. Dicer rescue assay in Dicer knockout embryonic stem cells using hDicer-specific helix mutants[10]**

(A) Experimental scheme of the Dicer rescue assay.

(B) Sequences of hDicer-specific helix mutants. Residues comprising the hDicer-specific helix are colored in blue. Mutant residues are in red.

(C) Results of TaqMan miRNA quantitative PCR using miR-16 and miR-21 probes (left). The plasmids were transfected in two different concentrations (20 mg and 30 mg). Western blot shows comparable expression of Dicer (right).

[10] Ibid.


A
Dicer expression plasmids
wild-type
Dicer specific helix mutants
transfection
Dicer-null mES
miRNA detection using Taqman qPCR
B
990 1000 1010 1020 1030
WT : HTSSRLNLL TPRHLNQKGK ALPLSSAEKR KAKWESLQNK QILVPELCAI
Point mutation : HTSSRLNLL TPRHLNQKGK ALPLSTGDKR KPKGDCGTNK QILVPELCAI
Deletion : HTSSRLNLL TPRHLNQKGK ALPLS----- --------NK QILVPELCAI
C
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
miR-16
miR-21
KO
20 30 20 30 20 30 : Dicer plasmid (μg)
WT
Deletion
Point mutation
KO WT Deletion Point mutation
0 20 30 20 30 20 30: Dicer plasmid (μg)
Dicer
GAPDH

Notably, Tian et al. discovered unexpected helix structure that protrudes out from the surface of the protein in a knob-like manner (**Figure II-18D**, helix in red). As amino acid sequence composing this helix is not found in PAZ domain of Ago proteins, it is designated as "Dicer-specific helix". In contrast to the previous substrate binding model in which bound dsRNA is aligned in parallel to the surface of platform-PAZ-connector helix cassette, this oriented siRNA duplex at an angle (60º) away from the protein surface. Contact between this helix and the bound siRNA was mainly accomplished through stacking interaction between Trp1014 and terminal base, and hydrogen bonds between Ser1005, Ser1006 to adjacent phosphates of the bound duplex (**Figure II-19B**). To test whether this helix has any role in Dicer processing, I generated two Dicer constructs whose Dicer specific helix is disrupted by deletion or point mutation (**Figure II-20A** and **Figure II-20B**). When these mutants were introduced into Dicer knockout cells, the mutants were as competent as the wild-type Dicer protein in miRNA production (**Figure II-20C**). These results suggest that the Dicer-specific helix is unlikely to be actively involved in miRNA processing.

Next, I and Tian et al. searched for structural evidence which supports 5′ end recognizing activity of human Dicer. Initial attempt was disappointing, as 5′ phosphate of 12-mer siRNA bound to Dicer PAZ cassette was located outward due to Dicer-specific helix. However, Tian et al.

**Figure II-21. Relative location of the 5′- and 3′-pockets in the crystal structure of human Dicer 'platform-PAZ-connector helix' cassette bound to RNA duplex in phosphate-containing solution[11]**

(A) A ribbon view of the 5′- and 3′-pockets. This view highlights the well-conserved 3′-end binding pocket (Y926, R927, Y961, Y962 and Y966), as well as bound inorganic phosphate, which is anchored by basic residues (R778, R780, R811 and H982), thereby revealing the potential 5′-phosphate binding pocket. R986 and R993 are in a disordered part of the structure (dashed line) which is in the vicinity of R778.

(B) An electrostatic surface view of the same region as in a, emphasizing the ≈ 20 Å distance separating the 5′- and 3′-pockets (represented by dashed circles). This is the expected distance between the 5′ and 3′ ends of a 2 nt overhang structure.

[11] This experiment was performed by Yuan Tian

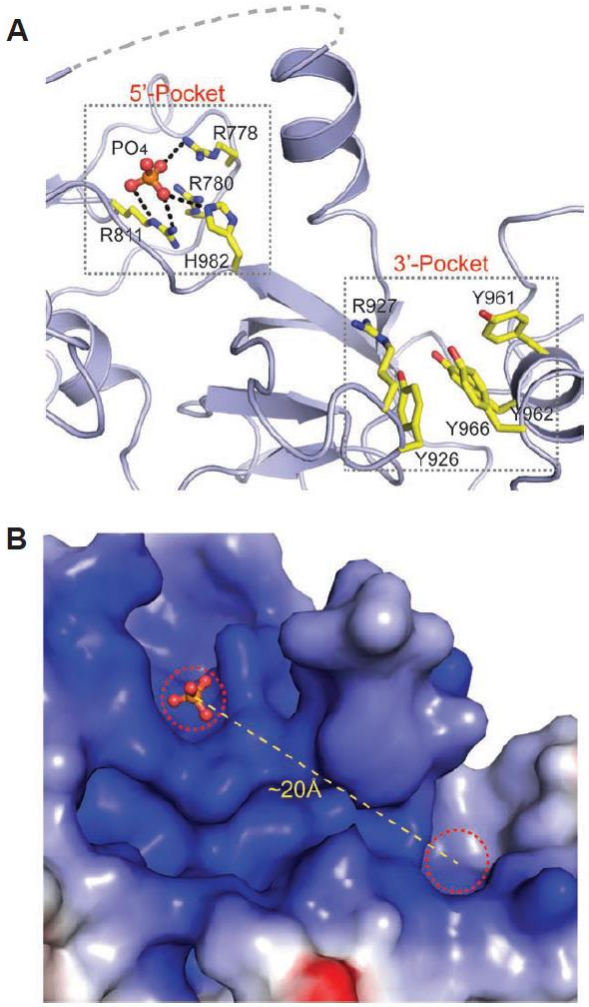

A
5'-Pocket
PO4
R778
R780
R811
H982
3'-Pocket
R927
Y961
Y962
Y966
Y926
B
~20Å

**Figure II-22. Structural details of the hDicer PAZ cassette bound to 16-mer siRNA and the impact of phosphate-pocket mutants in the hDicer PAZ cassette on binding affinity to a siRNA duplex[12]**

(A) Sequence of the self-complementary 16-mer siRNA containing 50-phosphate and UU overhangs at 3′ ends.

(B) Structure of the hDicer PAZ cassette bound to a 16-mer siRNA. The platform, PAZ, and connector helix are colored in yellow, blue, and pink, respectively. The Dicer PAZ-insertion element is disordered in this complex. The 5′-phosphate is highlighted in a space-filling representation.

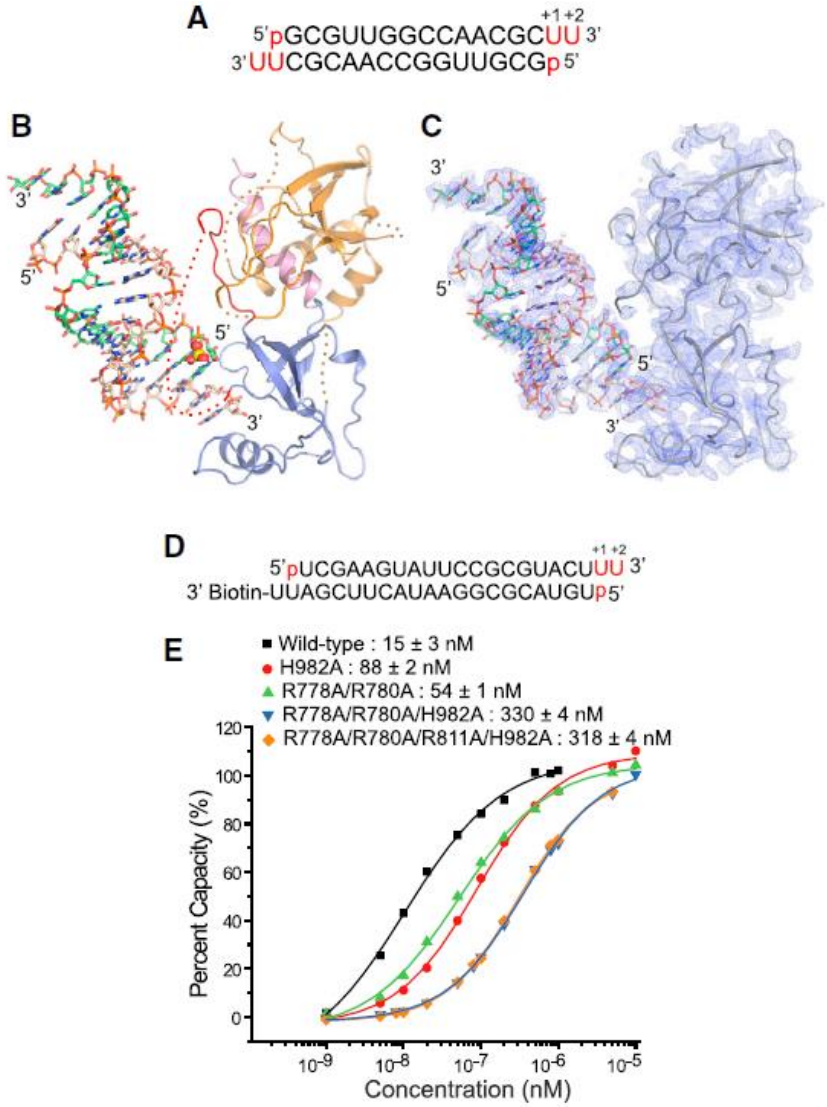
(C) Omit map (1s) of the protein and RNA in the hDicer PAZ cassette-16-mer siRNA complex.

(D) Sequence of immobilized RNA used for SPR measurements containing accessible 5′-phosphate, UU-3′ overhang and terminal U, U mismatch at one end and biotin attached to the 3′ overhang at the other end.

(E) Measurement of SPR-based RNA-binding affinities for hDicer PAZ cassette mutants located within the phosphate pocket

---

[12] Tian, Y., Simanshu, D.K., Ma, J.B., Park, J.E., Heo, I., Kim, V.N., and Patel, D.J. (2014). A phosphate-binding pocket within the platform-PAZ-connector helix cassette of human Dicer. Molecular cell *53*, 606-616.


A
+1 +2
5'pGCGUUGGCCAACGCUU 3'
3'UUCGCAACCGGUUGCGp 5'
B
3'
5'
5'
3'
C
3'
5'
5'
3'
D
+1 +2
5'pUCGAAGUAUUCCGCGUACUUU 3'
3' Biotin-UUAGCUUCAUAAGGCGCAUGUp5'
E
Wild-type : 15 ± 3 nM
H982A : 88 ± 2 nM
R778A/R780A : 54 ± 1 nM
R778A/R780A/H982A : 330 ± 4 nM
R778A/R780A/R811A/H982A : 318 ± 4 nM
Percent Capacity (%)
120
100
80
60
40
20
0
10-9
10-8
10-7
10-6
10-5
Concentration (nM)

discovered a sulfate ion which is coordinated to the side chains of Arg778, Arg780, Arg811, and His982 (**Figure II-21A**). Also, similar binding of inorganic phosphate to basic pocket was observed when 13-mer siRNA was used for structure. As residues comprising this basic pocket was identical to what has been found by biochemical study, this basic pocket is highly likely to be a structural counterpart of biochemically defined 5′-pocket. Supporting that conclusion, the bound phosphate is located ~20 Å away from the conserved 3′ end binding pocket (3′-pocket) in the PAZ domain, which is expected distance between the 5′ and 3′ ends of 2-nt overhang structure (**Figure II-21B**).

Moreover, Tian et al. found that Dicer PAZ cassette-16-mer siRNA structure in which Dicer-specific helix is melted and siRNA is positioned toward the surface of Dicer PAZ cassette, which implies that this helix structure can be dynamic and might not represent cleavage competent structure of Dicer (**Figure II-22**). This was consistent with the result that mutation or deletion of this helix does not disrupt Dicer processing in vitro.

## II-6. Role of Dicer helicase domain

As shown in **Figure II-8B**, interaction between 5′-pocket and 5′-end of RNA substrate was the most important determinant of Dicer processing site for ds-35 substrate. However, further study from other group revealed that pre-

miRNA loop is another important feature for Dicer substrate recognition and helicase domain is responsible for the recognition of loop structure (Castilla-Llorente et al., 2013; Flemr et al., 2013; Gu et al., 2012; Tsutsumi et al., 2011; Zhang and Zeng, 2010). To study the function of helicase domain in Dicer cleavage, I analyzed the published small RNA sequencing data which is performed using Dicer knockout mouse embryonic stem cell rescued with helicase deletion form of Dicer (Flemr et al., 2013). As helicase domain is also the site of TRBP interaction, I also used small RNA sequencing data from TRBP knockout cells to characterize the dependency on TRBP interaction. Also, I performed a rescue assay using Dicer construct which lacks ATPase activity of helicase domain (K70A).

Interestingly, deletion of helicase domain facilitated biogenesis of some miRNAs, while inhibiting others (**Figure II-23**). There was no correlation between miRNAs which are affected by helicase domain and 5′-pocket mutation, suggesting that each domain preferentially recognizes distinct subset of miRNAs. Similarly to 5′-pocket mutant, Dicer cleavage site was also affected in helicase domain deletion mutant, consistently in two replicates (**Figure II-24A**). Meanwhile, ATPase mutant construct showed no change in miRNA level and Dicer cleavage site (**Figure II-24A**), which suggests that although helicase domain is important for processing for subset of pre-miRNAs, ATPase activity is not required for this activity. Consistent with previous report (Fukunaga et al., 2012; Kim et al., 2014; Wilson et al.,

2015), TRBP knockout affected Dicer cleavage of only subset of miRNAs (**Figure II-24A**), so that the effect of helicase deletion mutant could not be explained by loss of interaction with TRBP. Comparison to 5′-pocket mutant data revealed that Dicer cleavage site change in helicase domain deletion is distinct from that in 5′-pocket mutant (**Figure II-24B**), which is consistent with idea that each domain has distinct preference for substrates. Together, helicase domain contributes to substrate recognition in addition to dsRNA end binding pockets, with different substrate preference. ATPase activity and TRBP binding does not explain this activity, suggesting that helicase domain of human Dicer protein have adopted unique activity for the recognition of pre-miRNA substrates.

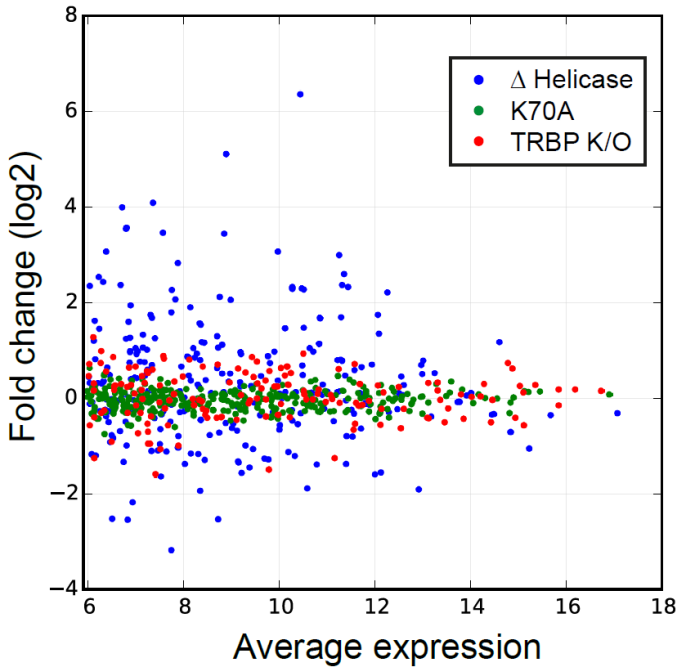



**Figure II-23. MiRNA expression profile of helicase domain mutants rescue sample and TRBP knockout cell line**

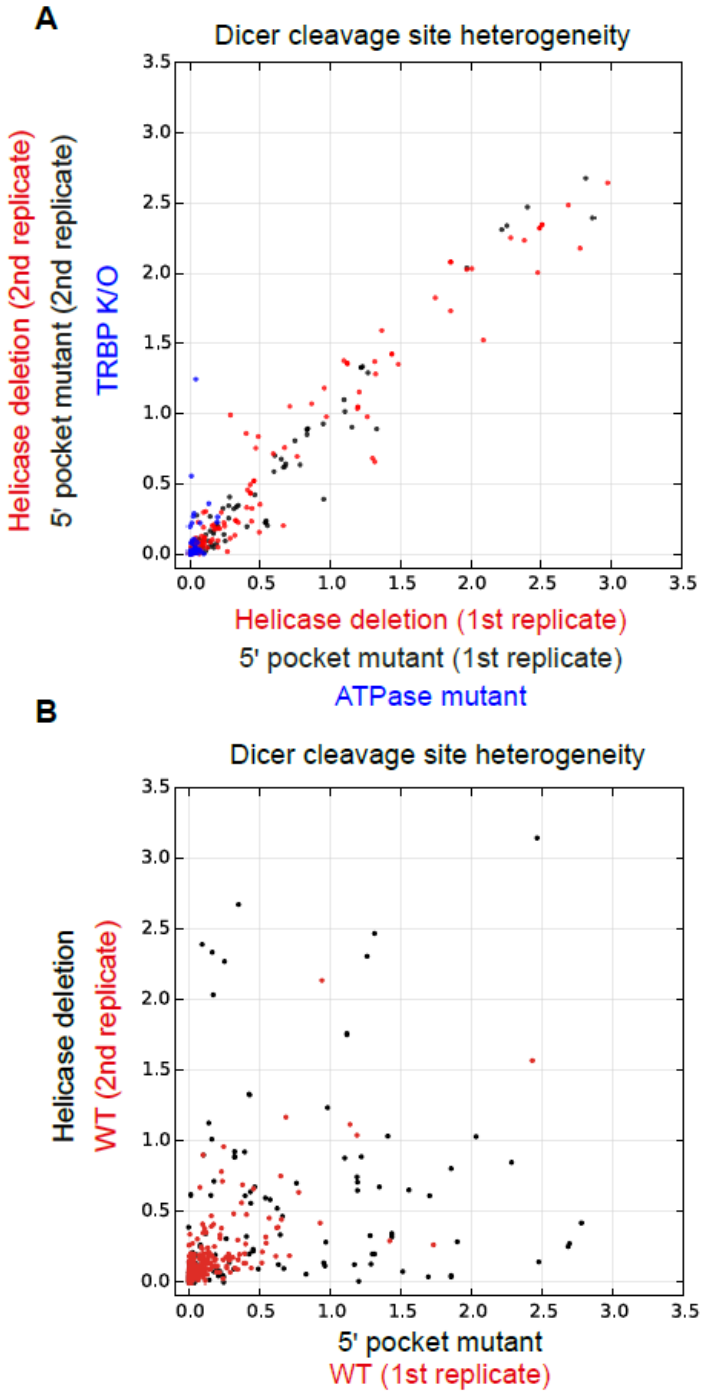
MA plot for miRNA expression pattern in Dicer Helicase deletion mutant rescue dataset[13] was compared to that of ATPase mutant (K70A) or TRBP knockout cell line.

[13] Flemr, M., Malik, R., Franke, V., Nejepinska, J., Sedlacek, R., Vlahovicek, K., and Svoboda, P. (2013). A retrotransposon-driven dicer isoform directs endogenous small interfering RNA production in mouse oocytes. Cell 155, 807-816.

**Figure II-24. Dicer cleavage site analysis of helicase mutants**

(A) Dicer cleavage site heterogeneity is measured by calculating the difference between Dicer cleavage patterns of mutant and control libraries. For helicase deletion mutant or 5′-pocket mutant, heterogeneity calculated using two replicates were plotted to show reproducibility. TRBP knockout and ATPase mutant dataset was plotted for comparison

(B) Dicer cleavage site heterogeneity of Helicase deletion mutant was compared to that of 5′-pocket mutant. As a control for natural variation in Dicer cleavage patterns, I calculated Dicer cleavage site heterogeneity between WT samples.


A
Dicer cleavage site heterogeneity
Helicase deletion (2nd replicate)
5' pocket mutant (2nd replicate)
TRBP K/O
Helicase deletion (1st replicate)
5' pocket mutant (1st replicate)
ATPase mutant
B
Dicer cleavage site heterogeneity
Helicase deletion
WT (2nd replicate)
5' pocket mutant
WT (1st replicate)

# CHAPTER III

# Conclusion

Biogenesis of miRNA is dependent on two successive cleavage of RNase III enzymes, Drosha and Dicer. As miRNAs search for their target using sequence complementarity, accurate target recognition and cleavage of Drosha and Dicer is critical in miRNA pathway.

In this study, I focused on the substrate recognition mechanism by Dicer. There are two tasks that Dicer should accomplish during the substrate recognition. First, Dicer should discriminate pre-miRNA from other similar structured RNAs, such as tRNAs, rRNAs and even hairpins in mRNAs. If Dicer non-specifically cleaves duplex RNAs, this would cause not only the degradation of functional RNAs but also the unwanted targeting of mRNAs which bears complementary sequences to the non-specific Dicer products. Meanwhile, Dicer should also have some flexibility in substrate recognition to accommodate pre-miRNAs which have diverse structural variations. Moreover, identification of terminal modification of pre-miRNAs by TUTases added structural complexity to pre-miRNA substrates.

Before I started this study, a proposed model for pre-miRNA recognition by Dicer emphasized the role of PAZ domain in recognition of 3′ end overhang of pre-miRNAs. However, according to this model, frequent modification at 3′ end of pre-miRNA would result in shift of Dicer cleavage site, resulting in mature miRNA with changed targeting potential. This

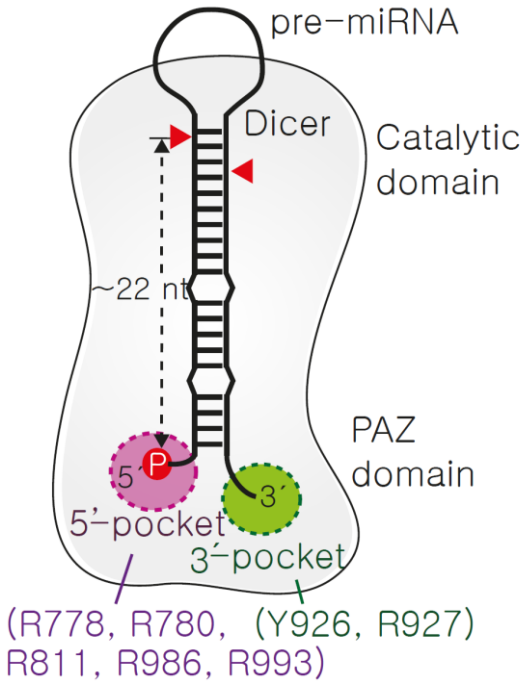



**Figure III-1. Model for Dicer substrate recognition**

could be either deleterious to miRNA pathway or advantageous, by providing novel regulatory mechanism. In this study, I tested the impact of 3′ end modification on pre-miRNA and discovered that contradictory to the previous model, human Dicer does not count from 3′ end to locate the cleavage site. Instead, it utilized 5′ end for cleavage site selection, which ensures consistent cleavage in spite of the frequent 3′ modification. 5′ end phosphate facilitated the recognition of the 5′ end. This phenomenon was conserved in Drosophila Dicer-1 but not in Giardia Dicer. Using amino acid conservation pattern and structural prediction, I could identify a group of basic residues which are responsible for the 5′ end recognition by Dicer, which I designated as 5′-pocket. These residues are conserved only in animal Dicer-1 family that function in miRNA production. Interestingly, only animal miRNAs, which is major substrate of animal Dicer-1, are known to utilize seed sequences (2-8nt relative to the 5′ end of miRNA) to search targets. As seed sequence can be dramatically changed by just 1 nt shift of 5′ end of miRNA, 5′ homogeneity is a critical feature of animal miRNAs. Thus I suggests that this 5′ end recognition mechanism also evolved to ensure the robust cleavage site selection which results in homogenous 5′ end of miRNA in spite of heterogeneously modified 3′ end of pre-miRNA (**Figure III-1**).

In collaboration with Dr. Dinshaw Patel group, I searched for the structural counterpart of 5′-pocket. For this end, Tian et al. solved series of

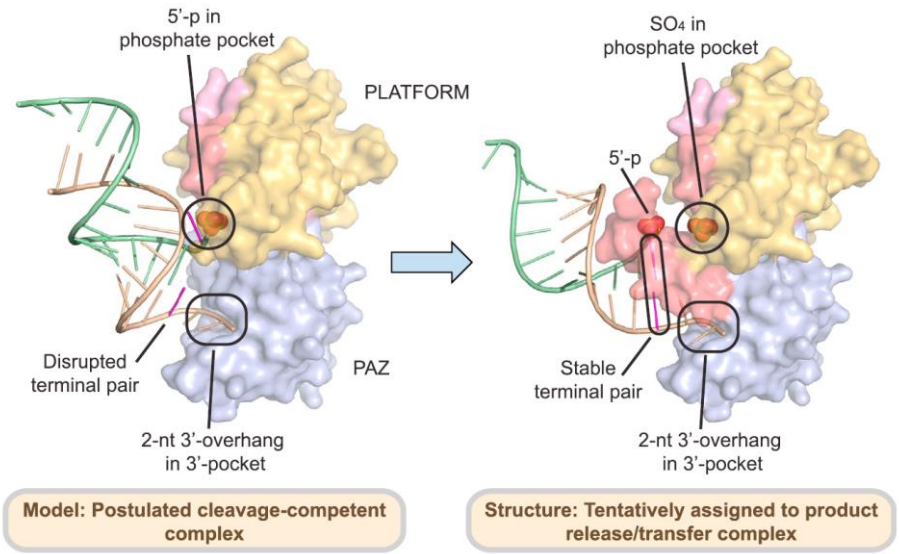



**Figure III-2. Proposed alignments of the hDicer PAZ cassette and siRNA in a postulated model of a cleavage-competent complex and in the structure tentatively assigned to the product release/transfer complex**

crystal structure of Dicer PAZ cassette and bound siRNA duplex. In most of this structure, 5′ end of pre-miRNA was located away from the surface of PAZ cassette. However, residues which are shown to participate in 5′-end recognition by my mutagenesis study indeed constituted a basic pocket which is occupied by inorganic phosphate or sulfate. Binding of 5′ end of siRNA duplex to this basic pocket was hindered by Dicer specific helix structure which extends out from the surface of PAZ cassette. Dicer mutated in the Dicer specific helix had no defect in miRNA production, suggesting that this helix structure is not involved in cleavage-competent binding of substrate. Moreover, in crystal structure of Dicer PAZ cassette and 16-mer siRNA duplex, this helix was melted/disordered and 5′ end phosphate located in proximity to the basic pocket and siRNA duplex was positioned in parallel with PAZ surface. Together, evidences suggest that the basic pocket which accommodates inorganic phosphate is likely to be the 5′-pocket. I propose that in cleavage competent binding, both 5′ end and 3′ end are captured by 5′-pocket and 3′-pocket in PAZ domain respectively, while I assume that formation of Dicer specific helix might represent conformation associated with product release/transfer (**Figure III-2**).

RNAi in mammalian systems is commonly induced by expressing small hairpin RNAs (shRNAs) from an RNA polymerase II or III promoter but the technology often suffers from inefficient and inaccurate Dicer processing (Chang et al., 2006; Kim and Rossi, 2007; Silva et al., 2004).

Based on our findings, a hairpin with a 5′-terminal phosphate and a 2 nt 3′ overhang should fit most optimally into the 5′-pocket and 3′-pocket of Dicer simultaneously. Also, it would be interesting to test whether a 5′-tri-phosphate could be efficiently accommodated into the 5′-pocket, as shRNAs driven by RNA polymerase III promoters bear a 5′-tri-phosphate. Understanding how human Dicer generates miRNAs will enable us to further improve the efficacy and safety of RNAi technology.

Although this study focused on the terminal recognition mechanism by Dicer, other features such as the shape of loop and stem length could be recognized by Dicer for proper substrate selection. For example, it had been reported that position of terminal loop can affect the heterogeneity of Dicer cleavage (Gu et al., 2012). Also, helicase domain of human or Drosophila Dicer-1 protein has been suggested to confer specificity on Dicer activity (Flemr et al., 2013; Gu et al., 2012; Ma et al., 2008; Tsutsumi et al., 2011). However how helicase domain recognizes loop structure of pre-miRNA substrate is still unknown. Crystal structure of helicase domain and pre-miRNA complex will provide answer to this question. Interestingly, although major human Dicer-1 substrate is pre-miRNA which is processed without ATP hydrolysis, amino acid sequence responsible for ATP hydrolysis are still retained in human Dicer-1, raising question about the role of ATP in human Dicer action mechanism. Finally, crystal structures of intact Dicer protein or Dicer in complex with its substrate or protein partners such as TRBP or Ago

have not been investigated yet. Together with an increasing number of studies on Dicer structure using cryo-EM technique (Lau et al., 2012; Taylor et al., 2013), this will provide us with comprehensive understanding on action mechanism of Dicer complex, including product release and transfer to Ago proteins.

# CHAPTER IV

# Materials and Methods

**Cell culture and transfection**

HEK293T cells were grown in DMEM (Welgene) supplemented with 10 % fetal bovine serum (Welgene). S2 cells were grown in HyClone SFX-Insect (Thermo Scientific) supplemented with 10 % fetal bovine serum (Welgene). Dicer knockout mouse embryonic stem (mES) cells (a gift from G. J. Hannon) were grown on mouse CF-1 feeder cells or gelatin-coated dishes in knockout DMEM (Gibco, Invitrogen) supplemented with 15 % fetal bovine serum (Gibco, Invitrogen), nonessential amino acids (Gibco, Invitrogen), 2 mM L-Glutamine (Sigma), 0.1 mM 2-mercaptoethanol (Sigma), and 1000 unit/ml leukemia inhibitory factor (Chemicon).

For HEK293T cells, transfection was carried out using the calcium-phosphate method. S2 cells were transfected using DDAB method as previously described (Han, 1996).

**Cloning and mutagenesis**

Giardia Dicer cDNA was amplified from Giardia Dicer-pFastBac HTa plasmid (a kind gift from J. Doudna) by PCR using the following primers: 5′-GGA TCC ATG CAT GCT TTG GGA CAC TG-3′ and 5′-GAT ATC GAG ACT GCA GGC TCT AGA TTC G-3′. PCR products were cloned into pGEM-T easy vector (Promega) and subsequently cloned into Flag-pcDNA3 vector (Invitrogen) at the BamHI and EcoRV site.

To introduce mutations into Dicer, QuickChange Site-Directed Mutagenesis Kit (Stratagene) was used. Mutated plasmids were confirmed by sequencing and subcloned into unmodified Flag-Dicer-pcDNA3.1 vector. The primer sequences used for the mutagenesis are provided in **Table IV-1**.

**Immunoprecipitation and in vitro Dicer processing**

For immunoprecipitation of Flag-Dicer, HEK293T cells were grown on 10-cm or 15-cm dishes and harvested at 48 hrs after Flag-Dicer-pcDNA3.1 expression plasmid transfection. The cells were incubated with lysis buffer (500 mM NaCl, 1 mM EDTA, 20 mM Tris (pH 8.0), 1 % Triton X-100) for 20 min on ice followed by sonication and centrifugation twice at 16,000 g for 10 min at 4 ℃. The supernatant was incubated with 10 ul of anti-Flag antibody conjugated to agarose beads (anti-Flag M2 affinity gel, Sigma) with constant rotation for 1 hr at 4 ℃. The beads were washed three times with lysis buffer and then four times with buffer D (200 mM KCl, 20 mM Tris (pH 8.0), 0.2 mM EDTA). The reactions were performed in a total volume of 30 μl in 2 mM MgCl2, 1 mM DTT, 1 unit/ul Ribonuclease inhibitor (Takara), 5′ end labeled pre-miRNA of 1x104 - 1x105 cpm and 15 ul of the immunopurified proteins in buffer D. The reaction mixture was incubated at 37 ℃ for 60~90 min. RNA was purified from the reaction mixture by phenol extraction and separated on 15 % urea polyacrylamide gel. Along with Decade marker (Ambion), synthetic hsa-let-7a RNA (22 nucleotides) was 5′-end labelled and

used as a size marker, because the 20-nucleotide RNA in Decade marker is often degraded to 18–19 nucleotides as we previously reported (Han et al., 2006).

For preparation of Drosophila Dicer-1, S2 cells confluent in 10-cm dish were transfected with Myc-Loquacious-PB-pRmHa3 expression plasmid (a kind gift from M. Siomi). After 1 day, 1 mM CuSO4 was added to the medium and cells were collected 2 days after CuSO4 treatment. The cells were incubated with lysis buffer for 30 min on ice, followed by sonication and centrifugation at 16,000 g for 10 min at 4 ℃. The supernatants were pre-cleared by incubation with 10 ul Protein A-Sepharose bead (GE Healthcare) for 2 hrs. Then, pre-cleared extract was incubated with 20 ul Protein A-Sepharose bead bound to anti-Myc antibody, 9E10, for 2 hrs at 4 ℃. The beads were washed three times with lysis buffer and then four times with buffer D and used for in vitro Dicer processing.

**Preparation of substrates**

Pre-let-7a-1, pre-miR-16-1, pre-miR-24-2 (mouse), pre-miR-142, pre-miR-143, pre-miR-200c and pre-miR-30a were synthesized by ST Pharm Co. Ltd. The sequences are presented in the figures. The pre-miRNA substrates with different 3′ overhang lengths and pre-miR-148b, pre-miR-27b, and pre-miR-151 were generated by ligating two synthetic single stranded RNAs as described previously. The sequences of RNA used for ligation are listed in

**Table IV-2**. The RNAs were labeled at the 5′ end with T4 polynucleotide kinase (T4 PNK, Takara) and [γ-$^{32}$P] ATP. Sequences of all endogenous pre-miRNAs used in our analysis are listed in **Table IV-3**.

For preparation of dsRNA substrates, a synthetic single stranded RNA was labeled at the 5′ end with [γ-$^{32}$P] ATP and T4 PNK. After phenol extraction, the labeled RNA was annealed to the complementary RNA by heating at 90 ℃ for 2 min and incubating at 30 ℃ for 2 hrs. In Fig. 1d, one strand of RNA was ligated to [α-$^{32}$P] pCp and treated with Calf Intestinal Alkaline Phosphatase (Takara) to generate the labeled 3′ end with a hydroxyl group. To attach a phosphate group at the 5′ end, 3′ end labeled RNAs were incubated with cold ATP and T4 PNK (Takara). Phenol extraction of RNA was performed after each reaction. Then, the labeled RNA was annealed to the RNA as described above.

**In vitro addition of uridine residues to pre-miRNAs**

A terminal nucleotidyl transferase, TUT4, is able to add 1-3 nt of uridine residues at the 3′ end of pre-miRNA in the absence of Lin28 protein in vitro (Heo et al., unpublished data). For this reaction, Flag-TUT4 expression plasmid was transfected in HEK293T cells using the calcium-phosphate method. After 48 hrs, total cell extract was prepared in buffer D by sonication and centrifugation at 16,000 g for 10 min at 4 ℃. 30 ul of reaction mixture contains 15 ul total cell extract (10 ug), 3.2 mM MgCl2, 1 mM DTT, 0.25

mM UTP, 1 unit/ul Ribonuclease inhibitor (Ambion), and 5′ end labeled pre-miRNA of 1x104 - 1x105 cpm. The reaction mixture was incubated at 37 ℃ for 15 min. After phenol extraction, the uridylated pre-miRNAs were gel purified and used for in vitro Dicer processing.

**3D structure prediction of Dicer fragment**

The 3D structure of Dicer fragment containing the PAZ domain was predicted by I-TASSER simulation (http://zhanglab.ccmb.med.umich.edu/I-TASSER) with amino acid sequences 751-1070. Crystal structure of PAZ domains from Dicer (PDB ID: 2ffl) and Argonautes (PDB ID: 1u04, 3dlb, 1r4k), together with RumA, a 23S ribosomal RNA methyltransferase (PDB ID: 1uwv), were used as templates for the comparative modeling. Among the five models predicted from the server, the one with a high C-score (-2.36) and an organized structure was chosen.

**Structure-based identification of the 5′-phosphate-binding pocket in human Dicer**

Diffraction quality crystals were grown for the complex of Dicer 'platform-PAZ-connector helix' cassette (residues 755 to 1055) and a self-complementary AGCGAAUUCGCUU duplex (underlined segment forms duplex) in phosphate-containing solution. The crystals of the complex belonged to space group I222, diffracted to 2.6 Å , and the structure of the

complex was refined to Rwork = 19.7 and Rfree = 23.7. In this structure, inorganic phosphate, which is anchored by basic residues (Arg778, Arg780, Arg811 and His982), reveals the potential 5′-phosphate binding pocket.

**Dicer rescue experiments**

For transfection, Dicer knockout mES cells were separated from feeder cells and 1,500,000 cells were seeded on gelatin-coated 6-well plates 1 day before transfection. 10 ug of plasmids (wild-type Dicer-pCK or 5′-mutant Dicer-pCK) were added to each well along with 10 ul of Lipofectamine 2000, according to the manufacturer's protocol (Invitrogen). Protein and RNA was extracted at 48hr post-transfection. To determine the protein levels, western blotting was performed using anti-Dicer and anti-tubulin (Abcam) antibodies. Expression of RNA was confirmed by northern blotting using the following probes: mmu-miR-293 (5′-ACA CTA CAA ACT CTG CGG CAC T-3′), mmu-miR-101a (5′-TTC AGT TAT CAC AGT ACT GTA-3′), mmu-miR-16 (5′-CGC CAA TAT TTA CGT GCT GCT A-3′) and tRNA-Lys-AAG (5′-GAG ATT AAG AGT CTC ATG CTC-3′).

To prepare small RNA cDNA libraries, RNA was extracted using TRIzol reagent (Invitrogen) or mirVana miRNA isolation kit (Ambion) and separated on 15% urea-PAGE. RNA of 17-26nt in length was gel-purified and ligated to the 3′ adapter using truncated T4 RNA ligase2 (NEB) in ATP-free condition. Subsequently, the ligation product was gel-purified and ligated to

the 5′ adapter using T4 RNA ligase1 (NEB). The final ligation product was gel-purified and used for reverse transcription using SuperScript II (Invitrogen). The cDNA was PCR-amplified with Phusion DNA polymerase (NEB). The resulting libraries were sequenced using Illumina Genome Analyzer II.

**Sequence analysis**

The essential workflow for early sequence analysis was performed as previously described (Hafner et al., 2010) with few modifications. After removing sequence reads including very low quality bases (< 10 in phred quality), 3′ adapter sequence was trimmed from the reads using a 5′-free variant of Smith-Waterman algorithm (scoring parameters: 2 for match, -3 for mismatch, -3 for linear gap). Then, we dropped short ($\leq$ 17nt) or repetitive sequences (0.7 and 1.5 for mono- or dinucleotide entropy of each sequence). The filtered sequences were aligned to Illumina adapter and primer sequences using the BWA short-read aligner (Li and Durbin, 2009) with 4 of allowed maximum edit distance, then matched reads were removed from further analysis. In the same way, the remaining sequences were aligned to the mouse genome mm9 assembly, which is downloaded from the University of California at Santa Cruz (http://genome.cse.ucsc.edu/). Annotations for aligned regions were retrieved using in-house software from RefSeq,

RepeatMasker and miRBase (downloaded from UCSC or miRBase on February 22, 2011). Software used in data processing and analysis can be downloaded from http://narrykim.org/s/park-dicer-2011/.

**Analysis of cleavage site change**

I first selected miRBase stem-loops that are relatively unaffected by reads aligned to multiple miRNA loci to avoid artifacts from over- or underestimated read counts. Stem-loops with more than 90% reads aligned to single stem-loop in every single sequencing lane were chosen for the later steps. Kullback-Leibler divergence (KLD) was used to quantify cleavage site change (difference of 5′ end position frequency) between two sequencing samples. To measure statistical significance of cleavage site change, Student's t-test was performed for KLDs between wild-type and mutant Dicer rescued samples, and KLDs between wild-type Dicer rescued samples and J1 (Babiarz et al., 2008), mouse embryo at 7.5 day (Chiang et al., 2010) or R1. Multiple testing correction was applied using the Benjamini-Hochberg method (Benjamini and Hochberg, 1995).

**Table IV-1. Primers used for mutagenesis**

| | | |
|---|---|---|
| R778A/R780A | forward | 5′- CTG ATG AAC TCA ACT TTG CAA GGG CGA AGC TCT ATC CTC C -3′ |
| | reverse | 5′- GGA GGA TAG AGC TTC GCC CTT GCA AAG TTG AGT TCA TCA G -3′ |
| R811A | forward | 5′- CAC TTT CCT GTG TAC ACA GCC TCT GGA GAG GTT ACC -3′ |
| | reverse | 5′- GGT AAC CTC TCC AGA GGC TGT GTA CAC AGG AAA GTG -3′ |
| R986A | forward | 5′- GTG GAC CAC ACA TCT TCA GCA CTT AAT CTT TTG ACA CCT C -3′ |
| | reverse | 5′- GAG GTG TCA AAA GAT TAA GTG CTG AAG ATG TGT GGT CCA C -3′ |
| H982A on R986A | forward | 5′- GCT GGA TGT GGA CGC CAC ATC TTC AGC AC -3′ |
| | reverse | 5′- GTG CTG AAG ATG TGG CGT CCA CAT CCA GC -3′ |
| R993A on R986A | forward | 5′- CAC TTA ATC TTT TGA CAC CTG CAC ATT TGA ATC AGA AGG -3′ |
| | reverse | 5′- CCT TCT GAT TCA AAT GTG CAG GTG TCA AAA GAT TAA GTG -3′ |
| R778A/R780A/R811A H982A/R986A/R993A (5′-mutant) | Subcloned H982A/R986A/R993A into R778A/R780A/R811A at the EcoRV and BamHI sites | |
| Y926F | forward | 5′- CGT TAT CAT TCC AAG ATT TCG CAA TTT TGA TCA GCC TC |

| | | |
|---|---|---|
| | | -3′ |
| | reverse | 5′- GAG GCT GAT CAA AAT TGC GAA ATC TTG GAA TGA TAA CG -3′ |
| R927A on Y926F (3′-mutant) | forward | 5′- CGT TAT CAT TCC AAG ATT TGC CAA TTT TGA TCA GCC TC -3′ |
| | reverse | 5′- GAG GCT GAT CAA AAT TGG CAA ATC TTG GAA TGA TAA CG -3′ |

**Table IV-2. RNA sequences used for generation of pre-miRNAs**

| | | |
|---|---|---|
| pre-let-7a-1 variants | common acceptor | 5′- UGA GGU AGU AGG UUG UAU AGU UUU AGG GUC ACA CC -3′ |
| | donor (+3) | 5′- CAC CAC UGG GAG AUA ACU AUA CAA UCU ACU GUC UUU CU -3′ |
| | donor (+4) | 5′- CAC CAC UGG GAG AUA ACU AUA CAA UCU ACU GUC UUU CUU -3′ |
| | donor (+5) | 5′- CAC CAC UGG GAG AUA ACU AUA CAA UCU ACU GUC UUU CUU U -3′ |
| pre-miR-16-1 variants | common acceptor | 5′- UAG CAG CAC GUA AAU AUU GGC GUU AAG AUU CU -3′ |
| | donor (+1) | 5′- AAA AUU AUC UCC AGU AUU AAC UGU GCU GCU GA -3′ |
| | donor (+3) | 5′- AAA AUU AUC UCC AGU AUU AAC UGU GCU GCU GAA U -3′ |
| pre-miR-142 variants | common acceptor | 5′- CAU AAA GUA GAA AGC ACU ACU AAC AGC ACU GGA G -3′ |
| | donor (+3) | 5′- GGU GUA GUG UUU CCU ACU UUA UGG AU -3′ |
| | donor (+4) | 5′- GGU GUA GUG UUU CCU ACU UUA UGG AUU -3′ |
| pre-miR-200c | common acceptor | 5′- CGU CUU ACC CAG CAG UGU UUG GGU GCG GUU -3′ |
| | donor (+3) | 5′- GGG AGU CUC UAA UAC UGC CGG GUA AUG AUG GA U -3′ |
| | donor (+4) | 5′- GGG AGU CUC UAA UAC UGC CGG GUA AUG AUG GA UU -3′ |

| | | |
|---|---|---|
| pre-miR-24-2 (mmu) variants | common acceptor | 5′- GUG CCU ACU GAG CUG AAA CAG UUG AUU CCA GUG CAC -3′ |
| | donor (+2) | 5′- UGG CUC AGU UCA GCA GGA ACA G -3′ |
| | donor (+3) | 5′- UGG CUC AGU UCA GCA GGA ACA GU -3′ |
| | donor (+4) | 5′- UGG CUC AGU UCA GCA GGA ACA GUU -3′ |
| | donor (+5) | 5′- UGG CUC AGU UCA GCA GGA ACA GUU U -3′ |
| pre-miR-148b | acceptor | 5′- GAA GUU CUG UUA UAC ACU CAG GCU GUG GCU CUC -3′ |
| | donor | 5′- UGA AAG UCA GUG CAU CAC AGA ACU UUG U -3′ |
| pre-miR-27b | acceptor | 5′- AGA GCU UAG CUG AUU GGU GAA CAG UGA UUG GUU -3′ |
| | donor | 5′- UCC GCU UUG UUC ACA GUG GCU AAG UUC UGC -3′ |
| pre-miR-151 | acceptor | 5′- UCG AGG AGC UCA CAG UCU AGU AUG UCU CAU -3′ |
| | donor | 5′- CCC CUA CUA GAC UGA AGC UCC UUG AGG -3′ |

**Table IV-3. Sequences of pre-miRNAs**

| pre-miRNA | Sequences (5′ to 3′) | Strand of mainly generated mature miRNA |
|---|---|---|
| pre-let-7a-1 | UGAGGUAGUAGGUUGUAUAGUUuuagggucacacccaccacugggagauaaCUAUACAAUCUACUGUCUUUC | 5p strand |
| pre-miR-16-1 | UAGCAGCACGUAAAUAUUGGCGuuaagauucuaaaauuaucuCCAGUAUUAACUGUGCUGCUG | 5p strand |
| pre-miR-143 | GGUGCAGUGCUGCAUCUCUGGUcaguugggagucUGAGAUGAAGCACUGUAGCUC | 3p strand |
| pre-miR-148b | GAAGUUCUGUUAUACACUCAGGCuguggcucucugaaagUCAGUGCAUCACAGAACUUUGU | 3p strand |
| pre-miR-27b | AGAGCUUAGCUGAUUGGUGAACagugauugguuuccgcuuugUUCACAGUGGCUAAGUUCUGC | 3p strand |
| pre-miR-151 | UCGAGGAGCUCACAGUCUAGUaugucucaucccccuaCUAGACUGAAGCUCCUUGAGG | both |

| | | |
|---|---|---|
| pre-miR-200c | CGUCUUACCCAGCAGUGUUUGGgugcgguuggg agucucUAAUACUGCCGGGUAAUGAUGGA | 3p strand |
| pre-miR-24-2 (mmu) | GUGCCUACUGAGCUGAAACAGUugauuccagug cacUGGCUCAGUUCAGCAGGAACAG | 3p strand |
| pre-miR-142 | CAUAAAGUAGAAAGCACUACUaacagcacuggag ggUGUAGUGUUUCCUACUUUAUGGA | both |
| pre-miR-30a | UGUAAACAUCCUCGACUGGAAGcugugaagccac agauggCUUUCAGUCGGAUGUUUGCAGC | both |

# REFERENCES

Ambros, V. (2004). The functions of animal microRNAs. Nature *431*, 350-355.

Babiarz, J.E., Ruby, J.G., Wang, Y., Bartel, D.P., and Blelloch, R. (2008). Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. Genes & development *22*, 2773-2785.

Bartel, D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. Cell *116*, 281-297.

Bartel, D.P. (2009). MicroRNAs: target recognition and regulatory functions. Cell *136*, 215-233.

Baum, J., Papenfuss, A.T., Mair, G.R., Janse, C.J., Vlachou, D., Waters, A.P., Cowman, A.F., Crabb, B.S., and de Koning-Ward, T.F. (2009). Molecular genetics and comparative genomics reveal RNAi is not functional in malaria parasites. Nucleic acids research *37*, 3788-3798.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. J Roy Stat Soc B Met *57*, 289-300.

Blaszczyk, J., Tropea, J.E., Bubunenko, M., Routzahn, K.M., Waugh, D.S., Court, D.L., and Ji, X. (2001). Crystallographic and modeling studies of RNase III suggest a mechanism for double-stranded RNA cleavage. Structure *9*, 1225-1236.

Bohnsack, M.T., Czaplinski, K., and Gorlich, D. (2004). Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. Rna *10*, 185-191.

Borsani, O., Zhu, J., Verslues, P.E., Sunkar, R., and Zhu, J.K. (2005). Endogenous siRNAs derived from a pair of natural cis-antisense

transcripts regulate salt tolerance in Arabidopsis. Cell *123*, 1279-1291.
Burroughs, A.M., Ando, Y., de Hoon, M.J., Tomaru, Y., Nishibu, T., Ukekawa, R., Funakoshi, T., Kurokawa, T., Suzuki, H., Hayashizaki, Y., *et al.* (2010). A comprehensive survey of 3' animal miRNA modification events and a possible role for 3' adenylation in modulating miRNA targeting effectiveness. Genome research *20*, 1398-1410.
Castilla-Llorente, V., Nicastro, G., and Ramos, A. (2013). Terminal loop-mediated regulation of miRNA biogenesis: selectivity and mechanisms. Biochemical Society transactions *41*, 861-865.
Cenik, E.S., Fukunaga, R., Lu, G., Dutcher, R., Wang, Y., Tanaka Hall, T.M., and Zamore, P.D. (2011). Phosphate and R2D2 restrict the substrate specificity of Dicer-2, an ATP-driven ribonuclease. Molecular cell *42*, 172-184.
Cerutti, H., and Casas-Mollano, J.A. (2006). On the origin and functions of RNA-mediated silencing: from protists to man. Current genetics *50*, 81-99.
Chang, K., Elledge, S.J., and Hannon, G.J. (2006). Lessons from Nature: microRNA-based shRNA libraries. Nature methods *3*, 707-714.
Chiang, H.R., Schoenfeld, L.W., Ruby, J.G., Auyeung, V.C., Spies, N., Baek, D., Johnston, W.K., Russ, C., Luo, S., Babiarz, J.E., *et al.* (2010). Mammalian microRNAs: experimental evaluation of novel and previously annotated genes. Genes & development *24*, 992-1009.
Colmenares, S.U., Buker, S.M., Buhler, M., Dlakic, M., and Moazed, D. (2007). Coupling of double-stranded RNA synthesis and siRNA generation in fission yeast RNAi. Molecular cell *27*, 449-461.
Denli, A.M., Tops, B.B., Plasterk, R.H., Ketting, R.F., and Hannon, G.J. (2004). Processing of primary microRNAs by the Microprocessor complex. Nature *432*, 231-235.
Drinnenberg, I.A., Weinberg, D.E., Xie, K.T., Mower, J.P., Wolfe, K.H., Fink, G.R., and Bartel, D.P. (2009). RNAi in budding yeast. Science

*326*, 544-550.
Elbashir, S.M., Lendeckel, W., and Tuschl, T. (2001). RNA interference is mediated by 21- and 22-nucleotide RNAs. Genes & development *15*, 188-200.
Flemr, M., Malik, R., Franke, V., Nejepinska, J., Sedlacek, R., Vlahovicek, K., and Svoboda, P. (2013). A retrotransposon-driven dicer isoform directs endogenous small interfering RNA production in mouse oocytes. Cell *155*, 807-816.
Forstemann, K., Tomari, Y., Du, T., Vagin, V.V., Denli, A.M., Bratu, D.P., Klattenhoff, C., Theurkauf, W.E., and Zamore, P.D. (2005). Normal microRNA maturation and germ-line stem cell maintenance requires Loquacious, a double-stranded RNA-binding domain protein. PLoS biology *3*, e236.
Fukunaga, R., Han, B.W., Hung, J.H., Xu, J., Weng, Z., and Zamore, P.D. (2012). Dicer partner proteins tune the length of mature miRNAs in flies and mammals. Cell *151*, 533-546.
Gasciolli, V., Mallory, A.C., Bartel, D.P., and Vaucheret, H. (2005). Partially redundant functions of Arabidopsis DICER-like enzymes and a role for DCL4 in producing trans-acting siRNAs. Current biology : CB *15*, 1494-1500.
Gregory, R.I., Yan, K.P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N., and Shiekhattar, R. (2004). The Microprocessor complex mediates the genesis of microRNAs. Nature *432*, 235-240.
Grishok, A., Pasquinelli, A.E., Conte, D., Li, N., Parrish, S., Ha, I., Baillie, D.L., Fire, A., Ruvkun, G., and Mello, C.C. (2001). Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control C. elegans developmental timing. Cell *106*, 23-34.
Gu, S., Jin, L., Zhang, Y., Huang, Y., Zhang, F., Valdmanis, P.N., and Kay, M.A. (2012). The loop position of shRNAs and pre-miRNAs is

critical for the accuracy of dicer processing in vivo. Cell *151*, 900-911.
Gunther, T. (2006). Concentration, compartmentation and metabolic function of intracellular free Mg2+. Magnesium research : official organ of the International Society for the Development of Research on Magnesium *19*, 225-236.
Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, M., Jr., Jungkamp, A.C., Munschauer, M., *et al.* (2010). Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. Cell *141*, 129-141.
Hammond, S.M., Boettcher, S., Caudy, A.A., Kobayashi, R., and Hannon, G.J. (2001). Argonaute2, a link between genetic and biochemical analyses of RNAi. Science *293*, 1146-1150.
Han, J., Lee, Y., Yeom, K.H., Kim, Y.K., Jin, H., and Kim, V.N. (2004). The Drosha-DGCR8 complex in primary microRNA processing. Genes & development *18*, 3016-3027.
Han, J., Lee, Y., Yeom, K.H., Nam, J.W., Heo, I., Rhee, J.K., Sohn, S.Y., Cho, Y., Zhang, B.T., and Kim, V.N. (2006). Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. Cell *125*, 887-901.
Han, K. (1996). An efficient DDAB-mediated transfection of Drosophila S2 cells. Nucleic acids research *24*, 4362-4363.
Heo, I., Joo, C., Cho, J., Ha, M., Han, J., and Kim, V.N. (2008). Lin28 mediates the terminal uridylation of let-7 precursor MicroRNA. Molecular cell *32*, 276-284.
Hutvagner, G., McLachlan, J., Pasquinelli, A.E., Balint, E., Tuschl, T., and Zamore, P.D. (2001). A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. Science *293*, 834-838.
Jiang, F., Ye, X., Liu, X., Fincher, L., McKearin, D., and Liu, Q. (2005).

Dicer-1 and R3D1-L catalyze microRNA maturation in Drosophila. Genes & development *19*, 1674-1679.
Ketting, R.F., Fischer, S.E., Bernstein, E., Sijen, T., Hannon, G.J., and Plasterk, R.H. (2001). Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. Genes & development *15*, 2654-2659.
Kim, D.H., and Rossi, J.J. (2007). Strategies for silencing human disease using RNA interference. Nature reviews Genetics *8*, 173-184.
Kim, V.N., Han, J., and Siomi, M.C. (2009). Biogenesis of small RNAs in animals. Nature reviews Molecular cell biology *10*, 126-139.
Kim, Y., Yeo, J., Lee, J.H., Cho, J., Seo, D., Kim, J.S., and Kim, V.N. (2014). Deletion of human tarbp2 reveals cellular microRNA targets and cell-cycle function of TRBP. Cell reports *9*, 1061-1074.
Knight, S.W., and Bass, B.L. (2001). A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in Caenorhabditis elegans. Science *293*, 2269-2271.
Kurihara, Y., and Watanabe, Y. (2004). Arabidopsis micro-RNA biogenesis through Dicer-like 1 protein functions. Proceedings of the National Academy of Sciences of the United States of America *101*, 12753-12758.
Landthaler, M., Yalcin, A., and Tuschl, T. (2004). The human DiGeorge syndrome critical region gene 8 and Its D. melanogaster homolog are required for miRNA biogenesis. Current biology : CB *14*, 2162-2167.
Lau, P.W., Guiley, K.Z., De, N., Potter, C.S., Carragher, B., and MacRae, I.J. (2012). The molecular architecture of human Dicer. Nature structural & molecular biology *19*, 436-440.
Lee, R.C., Feinbaum, R.L., and Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell *75*, 843-854.
Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P.,

Radmark, O., Kim, S., *et al.* (2003). The nuclear RNase III Drosha initiates microRNA processing. Nature *425*, 415-419.
Lee, Y.S., Nakahara, K., Pham, J.W., Kim, K., He, Z., Sontheimer, E.J., and Carthew, R.W. (2004). Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. Cell *117*, 69-81.
Lewis, B.P., Burge, C.B., and Bartel, D.P. (2005). Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell *120*, 15-20.
Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754-1760.
Lingel, A., Simon, B., Izaurralde, E., and Sattler, M. (2004). Nucleic acid 3'-end recognition by the Argonaute2 PAZ domain. Nature structural & molecular biology *11*, 576-577.
Liu, Q., Rand, T.A., Kalidas, S., Du, F., Kim, H.E., Smith, D.P., and Wang, X. (2003). R2D2, a bridge between the initiation and effector steps of the Drosophila RNAi pathway. Science *301*, 1921-1925.
Lund, E., Guttinger, S., Calado, A., Dahlberg, J.E., and Kutay, U. (2004). Nuclear export of microRNA precursors. Science *303*, 95-98.
Ma, E., MacRae, I.J., Kirsch, J.F., and Doudna, J.A. (2008). Autoinhibition of human dicer by its internal helicase domain. Journal of molecular biology *380*, 237-243.
Ma, J.B., Ye, K., and Patel, D.J. (2004). Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. Nature *429*, 318-322.
MacRae, I.J., Zhou, K., and Doudna, J.A. (2007). Structural determinants of RNA recognition and cleavage by Dicer. Nature structural & molecular biology *14*, 934-940.
Macrae, I.J., Zhou, K., Li, F., Repic, A., Brooks, A.N., Cande, W.Z., Adams, P.D., and Doudna, J.A. (2006). Structural basis for double-stranded RNA processing by Dicer. Science *311*, 195-198.

Miyoshi, K., Miyoshi, T., Hartig, J.V., Siomi, H., and Siomi, M.C. (2010). Molecular mechanisms that funnel RNA precursors into endogenous small-interfering RNA and microRNA biogenesis pathways in Drosophila. Rna *16*, 506-515.
Mourelatos, Z., Dostie, J., Paushkin, S., Sharma, A., Charroux, B., Abel, L., Rappsilber, J., Mann, M., and Dreyfuss, G. (2002). miRNPs: a novel class of ribonucleoproteins containing numerous microRNAs. Genes & development *16*, 720-728.
Mukherjee, K., Campos, H., and Kolaczkowski, B. (2013). Evolution of animal and plant dicers: early parallel duplications and recurrent adaptation of antiviral RNA binding in plants. Molecular biology and evolution *30*, 627-641.
Murchison, E.P., Partridge, J.F., Tam, O.H., Cheloufi, S., and Hannon, G.J. (2005). Characterization of Dicer-deficient murine embryonic stem cells. Proceedings of the National Academy of Sciences of the United States of America *102*, 12135-12140.
Olsen, P.H., and Ambros, V. (1999). The lin-4 regulatory RNA controls developmental timing in Caenorhabditis elegans by blocking LIN-14 protein synthesis after the initiation of translation. Developmental biology *216*, 671-680.
Park, W., Li, J., Song, R., Messing, J., and Chen, X. (2002). CARPEL FACTORY, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in Arabidopsis thaliana. Current biology : CB *12*, 1484-1495.
Provost, P., Dishart, D., Doucet, J., Frendewey, D., Samuelsson, B., and Radmark, O. (2002). Ribonuclease activity and RNA binding of recombinant human Dicer. The EMBO journal *21*, 5864-5874.
Roy, A., Kucukural, A., and Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. Nature protocols *5*, 725-738.

Saito, K., Ishizuka, A., Siomi, H., and Siomi, M.C. (2005). Processing of pre-microRNAs by the Dicer-1-Loquacious complex in Drosophila cells. PLoS biology *3*, e235.
Serra, M.J., Baird, J.D., Dale, T., Fey, B.L., Retatagos, K., and Westhof, E. (2002). Effects of magnesium ions on the stabilization of RNA oligomers of defined structures. Rna *8*, 307-323.
Shabalina, S.A., and Koonin, E.V. (2008). Origins and evolution of eukaryotic RNA interference. Trends in ecology & evolution *23*, 578-587.
Silva, J., Chang, K., Hannon, G.J., and Rivas, F.V. (2004). RNA-interference-based functional genomics in mammalian cells: reverse genetics coming of age. Oncogene *23*, 8401-8409.
Song, J.J., Liu, J., Tolia, N.H., Schneiderman, J., Smith, S.K., Martienssen, R.A., Hannon, G.J., and Joshua-Tor, L. (2003). The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. Nature structural biology *10*, 1026-1032.
Tabara, H., Sarkissian, M., Kelly, W.G., Fleenor, J., Grishok, A., Timmons, L., Fire, A., and Mello, C.C. (1999). The rde-1 gene, RNA interference, and transposon silencing in C. elegans. Cell *99*, 123-132.
Taylor, D.W., Ma, E., Shigematsu, H., Cianfrocco, M.A., Noland, C.L., Nagayama, K., Nogales, E., Doudna, J.A., and Wang, H.W. (2013). Substrate-specific structural rearrangements of human Dicer. Nature structural & molecular biology *20*, 662-670.
Tian, Y., Simanshu, D.K., Ma, J.B., Park, J.E., Heo, I., Kim, V.N., and Patel, D.J. (2014). A phosphate-binding pocket within the platform-PAZ-connector helix cassette of human Dicer. Molecular cell *53*, 606-616.
Tomari, Y., Matranga, C., Haley, B., Martinez, N., and Zamore, P.D. (2004). A protein sensor for siRNA asymmetry. Science *306*, 1377-1380.

Tsutsumi, A., Kawamata, T., Izumi, N., Seitz, H., and Tomari, Y. (2011). Recognition of the pre-miRNA structure by Drosophila Dicer-1. Nature structural & molecular biology *18*, 1153-1158.
Ullu, E., Tschudi, C., and Chakraborty, T. (2004). RNA interference in protozoan parasites. Cellular microbiology *6*, 509-519.
Vazquez, F., Vaucheret, H., Rajagopalan, R., Lepers, C., Gasciolli, V., Mallory, A.C., Hilbert, J.L., Bartel, D.P., and Crete, P. (2004). Endogenous trans-acting siRNAs regulate the accumulation of Arabidopsis mRNAs. Molecular cell *16*, 69-79.
Vermeulen, A., Behlen, L., Reynolds, A., Wolfson, A., Marshall, W.S., Karpilow, J., and Khvorova, A. (2005). The contributions of dsRNA structure to Dicer specificity and efficiency. Rna *11*, 674-682.
Volpe, T.A., Kidner, C., Hall, I.M., Teng, G., Grewal, S.I., and Martienssen, R.A. (2002). Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. Science *297*, 1833-1837.
Welker, N.C., Maity, T.S., Ye, X., Aruscavage, P.J., Krauchuk, A.A., Liu, Q., and Bass, B.L. (2011). Dicer's helicase domain discriminates dsRNA termini to promote an altered reaction mode. Molecular cell *41*, 589-599.
Wilson, R.C., Tambe, A., Kidwell, M.A., Noland, C.L., Schneider, C.P., and Doudna, J.A. (2015). Dicer-TRBP complex formation ensures accurate mammalian microRNA biogenesis. Molecular cell *57*, 397-407.
Wu, H., Ye, C., Ramirez, D., and Manjunath, N. (2009). Alternative processing of primary microRNA transcripts by Drosha generates 5' end variation of mature microRNA. PloS one *4*, e7566.
Xie, Z., Allen, E., Wilken, A., and Carrington, J.C. (2005). DICER-LIKE 4 functions in trans-acting small interfering RNA biogenesis and vegetative phase change in Arabidopsis thaliana. Proceedings of the National Academy of Sciences of the United States of America *102*, 12984-12989.

Xie, Z., Johansen, L.K., Gustafson, A.M., Kasschau, K.D., Lellis, A.D., Zilberman, D., Jacobsen, S.E., and Carrington, J.C. (2004). Genetic and functional diversification of small RNA pathways in plants. PLoS biology *2*, E104.
Yan, K.S., Yan, S., Farooq, A., Han, A., Zeng, L., and Zhou, M.M. (2003). Structure and conserved RNA binding of the PAZ domain. Nature *426*, 468-474.
Yi, R., Qin, Y., Macara, I.G., and Cullen, B.R. (2003). Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. Genes & development *17*, 3011-3016.
Zhang, H., Kolb, F.A., Brondani, V., Billy, E., and Filipowicz, W. (2002). Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP. The EMBO journal *21*, 5875-5885.
Zhang, H., Kolb, F.A., Jaskiewicz, L., Westhof, E., and Filipowicz, W. (2004). Single processing center models for human Dicer and bacterial RNase III. Cell *118*, 57-68.
Zhang, X., and Zeng, Y. (2010). The terminal loop region controls microRNA processing by Drosha and Dicer. Nucleic acids research *38*, 7689-7697.

# PUBLICATIONS

**J.-E. Park*, I. Heo*, Y. Tian, D. K. Simanshu, H. Chang, D. Jee, D. J. Patel and V. N. Kim (2011) "Dicer recognizes the 5′ end of RNA for efficient and accurate processing" Nature, 475:201-205. (*co-first authors)**

**I. Heo, M. Ha, J. Lim, M.-J. Yoon, J.-E. Park, S. C. Kwon, H. Chang and**

**V. N. Kim (2012) "Mono-uridylation of pre-microRNA as a key step in the biogenesis of group II let-7 microRNAs" Cell, 151: 521-532.**

**Y. Tian, D. K. Simanshu, J.-B. Ma, J.-E. Park, I. Heo, V. N. Kim, and D. J. Patel (2014) "A Phosphate-Binding Pocket within the Platform-PAZ-Connector Helix Cassette of Human Dicer" Molecular Cell, 53:606-616.**

**Y. Kim, J. H. Lee, J.-E. Park, J. Cho, H. Yi and V. N. Kim (2014) "PKR is activated by cellular dsRNAs during mitosis and acts as a mitotic regulator" Genes & Development, 28: 1310-1322.**

**K. Boo, J. Bhin, Y. Jeon, J. Kim, H. J. Shin, J. -E. Park, K. Kim, C. R. Kim, H. Jang, I. H. Kim, V. N. Kim, D. Hwang, H. Lee, S. H. Baek (2015) "Pontin functions as an essential coactivator for Oct4-dependent lincRNA expression in mouse embryonic stem cells" Nature Communications, 6:6810.**

# 국문 초록

마이크로RNA는 타겟 유전자의 mRNA와 상보적인 염기 서열을 이용해 결합하여 다양한 유전자의 발현을 억제하는 역할을 한다. Small RNA 대용량 서열분석 결과에 의하면 이러한 마이크로RNA의 말단에 1~2개의 유리딘 잔기 또는 아데닌 잔기가 결합된 형태가 많이 존재한다. 따라서, 이러한 마이크로RNA의 말단 변이가 어떤 생물학적 의미를 가질 수 있는지에 대한 연구가 수행되고 있다.

마이크로RNA의 말단 변이는 활성화된 형태의 mature 마이크로RNA(mature miRNA)뿐 아니라, 그 전구체인 precursor 마이크로RNA (pre-miRNA)에서도 관찰된다. Dicer는 pre-miRNA를 절단하여 mature miRNA를 생산하는 효소로써, 기존의 연구 결과에 의하면 RNA 이중가닥의 3′ 말단 부위에 결합하여 이 부분으로부터 일정한 길이를 재서 이중가닥을 끊어 내는 “줄자와 가위” 역할을 하는 것으로 알려져 있었다. 이 모델에 의하면 pre-miRNA에 3′ 말단 변이가 일어난 경우 Dicer 절단 부위도 이에 따라 변하게 될 것이라고 예상할 수 있는데, 이렇게 되면 mature miRNA의 염기 서열 및 타겟 유전자 군이 뒤바뀌게 된다. 이러한 일이 실제로 일어나는지 확인하기 위해 나는 다양한 3′ 말단 변이를 가진 pre-miRNA들과 Dicer 효소를 섞어 주고, 생성물의 길이 및 양을 관찰하였다. 그 결과 예상과 달리 3′ 말단 변이와 관계 없이 Dicer가 같은 곳을 자를 수 있다는 것을 확인하였으며, Dicer가 3′ 말단 뿐만 아니라 5′ 말단을 함께 인지하며, 자르는 부위를 정하는 데에는 5′ 말단으로부터의 거리가 더 중요하다는 것을 밝혔다. 5′ 말단 인지는 pre-miRNA의 말단 구조가 풀리기

쉬운 구조일 경우에 더 잘 관찰되었으며, 실제 세포 내에서의 조건과 유사한 양의 마그네슘 이온을 사용하면 5′ 말단 인지가 3′ 말단 인지에 비해 훨씬 우세하게 작용한다는 것을 확인하였다. 따라서 세포 내에서는 5′ 말단 인지가 Dicer 절단 부위 결정에 가장 중요하게 작용할 것이라는 가설을 제시하였다. Dicer의 여러 돌연변이체를 검사하여 5′ 말단 인지 기능을 담당하는 아미노산 잔기를 찾아내었고, 이들이 구성하는 가상의 구조를 5′ -pocket으로 명명하였다. 이 5′ -pocket은 초파리의 Dicer-1에는 보존되어 있었지만 Giardia와 같은 원시 진핵생물의 Dicer에는 보존되어 있지 않았다.

5′ -pocket을 망가뜨린 Dicer 돌연변이를 Dicer가 knockout된 생쥐배아줄기세포에 발현시키는 실험을 이용하여, 나는 5′ -pocket이 마이크로RNA 생성의 효율성과 정확성을 위해 매우 중요하다는 것을 확인하였다. 또한 Dinshaw Patel 박사 그룹과의 공동 연구를 통해서, 5′ pocket에 해당하는 구조를 확인하고, Dicer의 pre-miRNA 인지 과정에 대한 종합적인 모델을 제시하였다. 본 연구 결과는 마이크로RNA 생성 과정에서 정확한 정보의 선별 및 전달이 어떻게 이루어지는지 제시함으로써 RNA 결합 단백질의 기질 선별 방법이 용도에 따라 어떻게 진화해 가는지에 대한 하나의 모델을 제시하였다. 또한, 유전자 치료제로 주목 받고 있는 short hairpin RNA의 정확도를 높이는 데 기여할 것이라 예상된다.